DOCUMENT RESUME

ED 210 504 CE 030 776

TITLE	Microcomputer Hardware. Energy Technology Series.
INSTITUTION	Technical Education Research Centre-Southwest, Waco, Tex.
SPONS AGENCY	Office of Vocational and Adult Education (ED), Washington, D.C.
BUREAU NO	498AH80027
PUB DATE	Dec 80
CONTRACT	300-78-0551
NOTE	216p.; For related documents see CE 030 771-789 and ED 190 746-761.
AVAILABLE FROM	Center for Occupational Research and Development, 601 Lake Air Dr., Waco, TX 76710 ($2.50 per module; $17.50 for entire course).
EDRS PRICE	MF01 Plus Postage. PC Not Available from EDRS.
DESCRIPTORS	Adult Education; *Analog Computers; Behavioral Objectives; Course Descriptions; Courses; *Digital Computers; *Energy; Energy Conservation; Glossaries; Input Output; Input Output Devices; Laboratory Experiments; Learning Activities; Learning Modules; *Microcomputers; Postsecondary Education; *Power Technology; *Technical Education; Two Year Colleges
IDENTIFIERS	Conversion (Format); Troubleshooting

ABSTRACT

This course in microcomputer hardware is one of 16 courses in the Energy Technology Series developed for an Energy Conservation-and-Use Technology curriculum. Intended for use in two-year postsecondary technical institutions to prepare technicians for employment, the courses are also useful in industry for updating employees in company-sponsored training programs. Comprised of seven modules, the course surveys integrated circuit logic, common electrical and logical digital interfacing techniques, techniques for getting digital and analog data into and out of microcomputers, applications of these techniques to actual control problems, data communication ideas, and microcomputer troubleshooting techniques. Written by a technical expert and approved by industry representatives, each module contains the following elements: introduction, prerequisites, objectives, subject matter, exercises, laboratory materials laboratory procedures (experiment section for hands-on portion), data tables (included in most basic courses to help students learn to collect or organize data), references, and glossary. Module titles are Digital Components, Semi-Conductor Logic Families, Input/Output Devices and Techniques, Analog/Digital Conversion, Data Communication, Bus Systems, and Troubleshooting Microcomputer Components. (YLB)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED210504

MICROCOMPUTER

HARDWARE

TECHNICAL EDUCATION RESEARCH CENTER - SW

4800 LAKEWOOD DRIVE - SUITE 5

WACO, TEXAS 76710

DEC. 1980

CE030776

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

✓ This document has been reproduced as received from the person or organization originating it.
Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official NIE position or policy.

"PERMISSION TO REPRODUCE THIS MATERIAL IN MICROFICHE ONLY HAS BEEN GRANTED BY

D. M. Hull

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

P R E F A C E

ABOUT ENERGY TECHNOLOGY MODULES

The modules were developed by TERC-SW for use in two-year postsecondary technical institutions to prepare technicians for employment and are useful in industry for updating employees in company-sponsored training programs. The principles, techniques and skills taught in the modules, based on tasks that energy technicians perform, were obtained from a nationwide advisory committee of employers of energy technicians. Each module was written by a technical expert and approved by representatives from industry.

A module contains the following elements:

Introduction, which identifies the topic and often includes a rationale for studying the material.

Prerequisites, which identify the material a student should be familiar with before studying the module.

Objectives, which clearly identify what the student is expected to know for satisfactory module completion. The objectives, stated in terms of action-oriented behaviors, include such action words as operate, measure, calculate, identify and define, rather than words with many interpretations, such as know, understand, learn and appreciate.

Subject Matter, which presents the background theory and techniques supportive to the objectives of the module. Subject matter is written with the technical student in mind.

Exercises, which provide practical problems to which the student can apply this new knowledge.

Laboratory Materials, which identify the equipment required to complete the laboratory procedure.

Laboratory Procedures, which is the experiment section, or "hands-on" portion, of the module (including step-by-step instruction) designed to reinforce student learning.

Data Tables, which are included in most modules for the first year (or basic) courses to help the student learn how to collect and organize data.

References, which are included as suggestions for supplementary reading/viewing for the student.

Test, which measures the student's achievement of the prestated objectives.

Glossary, which defines and explains terms or words used within the module that are uncommon, technical, or anticipated as being unfamiliar to the student.

CONTENTS

Preface

Table of Contents

MODULE MH-01 Digital Components

MODULE MH-02 Semi-Conductor Logic Families

MODULE MH-03 Input/Output Devices and Techniques

MODULE MH-04 Analog/Digital Conversion

MODULE MH-05 Data Communication

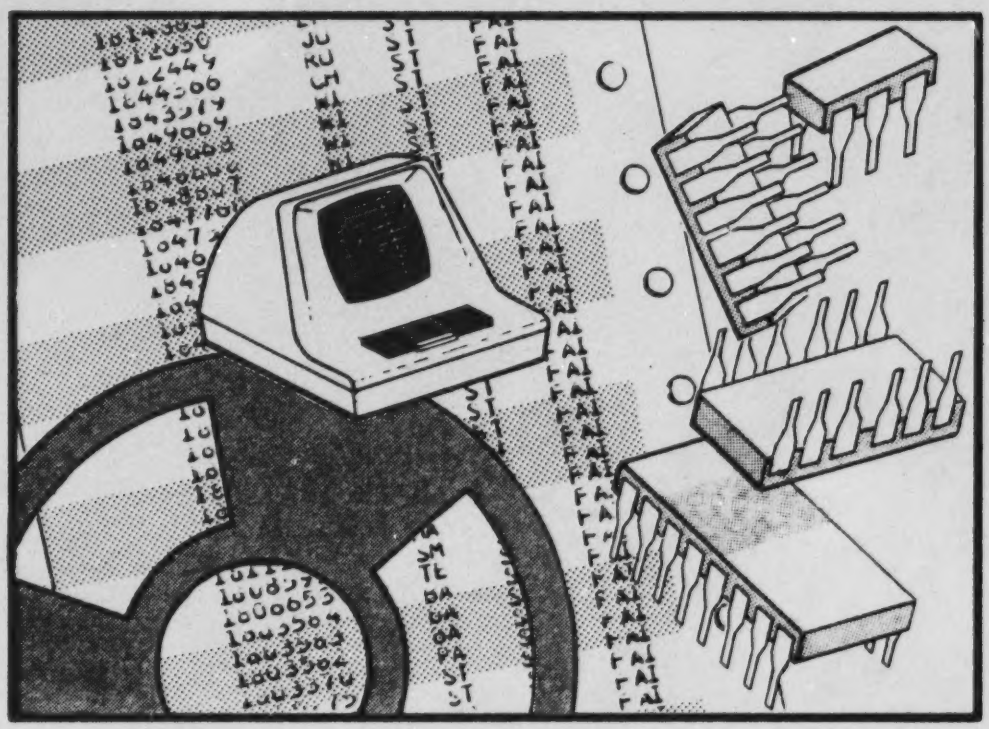
MODULE MH-06 Bus Systems

MODULE MH-07 Troubleshooting Micrcomputer Components

ENERGY TECHNOLOGY

CONSERVATION AND USE

MICROCOMPUTER HARDWARE



MODULE MH-01

DIGITAL COMPONENTS

TECHNICAL EDUCATION RESEARCH CENTER - SOUTHWEST
4800 LAKEWOOD DRIVE, SUITE 5
WACO, TEXAS 76710

Copyright, 1979 by the Technical Education Research Center - SW. Printed in the United States of America. All rights reserved. These materials, or parts thereof, may not be reproduced in any form without permission.

Copyright on these materials is claimed only during the period of development, test and evaluation, unless authorization is granted by the U. S. Office of Education to claim copyright also on the final materials. For information on the status of the copyright claim, contact either the copyright proprietor or the U. S. Office of Education.

The materials contained herein were developed pursuant to a contract with the Office of Education, U. S. Department of Health, Education and Welfare: Project No. 498AH0027, Contract No. 300 780551. Contractors undertaking such projects under government sponsorship are encouraged to express freely their professional judgement in the conduct of the project. Points of view or opinions stated do not, therefore, necessarily represent official Office of Education position or policy.

INTRODUCTION

Computers are logic devices composed of a vast array of elementary logic gates. These gates often are used singly or in simple combinations in computer support circuits such as memory, input/output and clocks. This module investigates the logical properties of these gates and some of the basic digital circuits that can be made from them.

PREREQUISITES

The student should have completed the course in Microcomputer Operations and have the ability to use d.c. instruments and to breadboard integrated circuits (IC).

OBJECTIVES

Upon completion of this module, the student should be able to:

1. Predict output from multiple gate circuit schematics by using truth tables and standard logic symbols.
2. Define the following terms:
   a. Microcomputer.
   b. Large scale integration (LSI).
   c. Hardware.
   d. Software.
   e. Arithmetic unit.
   f. Clock and control unit.
   g. Memory.
   h. Data.
   i. Instruction.

j. Central processing unit (CPU).
k. Primary memory.
l. Secondary memory.
m. Location.
n. Address.
o. Read only memory (ROM).
p. Random access memory (RAM).
q. Peripheral.
r. Digital signal.
s. Level signal.
t. Pulse signal.
u. Gate.
v. AND gate.
w. OR gate.
x. NOT gate (or inverter).
y. NAND gate.
z. NOR gate.
aa. Exclusive OR gate.
bb. Non-exclusive OR gate.
cc. Inversion bubble.
dd. Adder.
ee. Register.
ff. Buffer
gg. Flip-flop.
hh. Set-reset flip-flop.
ii. Trigger flip-flop.
jj. Master-slave flip-flop.
kk. J-K (steerable) flip-flop.
ll. Program counter.
mm. Leading edge.
nn. Trailing edge.

oo. Identity function.

pp. Serial parity.

CHANGES AND THE COMPUTER'S COMPONENTS

Since the first generation of digital computer hardware appeared in the 1940's, the momentum of changing computer technology has grown at an ever-increasing pace.

The challenge for technicians involved with computer hardware, including technicians in energy conservation and use, is to keep abreast of these rapidly-changing developments.

This introductory course in computer hardware summarizes the basic logic components found in all digital computers, as well as basic digital electronics and simple troubleshooting techniques. Students should realize, however, that in this environment of technological change, the hardware encountered in the field will be wide-ranging in appearance and function.

The term "microcomputer" refers to a digital computing device that utilizes an advanced solid-state microprocessor. Microprocessors were first introduced in 1971 as a result of Large Scale Integration (LSI) technology, which enabled the construction of a single silicon Integrated Circuit (IC) containing the miniaturized logic devices required by computers, which will be detailed in this course.

Hardware refers to the physical components and circuitry that make up a digital computer. Software refers to programs, or rules for organizing operations, by which the computer superimposes symbolic meaning upon interrelating binary electrical conditions.

The following is an overview of the four basic groups of components that make up a computer:

GROUP 1. ARITHMETIC UNIT. The unit that gives a computer its name, the arithmetic unit utilizes high-speed dig-

ital circuitry to compute basic arithmetic operations. All computer information processing is accomplished through symbolic numeric coding that undergoes simple mathematic transformations.

GROUP 2. CLOCK AND CONTROL UNIT. Using stored programs to guide complex functions sequentially through organized, logical routines, the clock and control unit provides the necessary timing and control signals to coordinate the many internal operations of a computer.

In combination, the arithmetic unit and the clock and control unit are sometimes referred to as the central processing unit (CPU).

GROUP 3. MEMORY. Computer memory can serve as a holding tank for partial answers while other parts of a problem are being solved. Information of this nature being stored for the arithmetic unit is commonly referred to as data. Memory also stores information for the control unit; this information is called instructions. Programs of routines and subroutines offer step-by-step guidance, enabling the computer to perform a task.

The term "primary" memory refers to internal storage made up of magnetic cores or semiconductors; "secondary" memory consists of external devices, such as magnetic tape or disk recorder/readers. Modern semiconductor memories have increased computer reliability and speed, making larger memory capacities economically feasible. New advances in magnetic-bubble memory promise to continue that trend.

Memory is divided into units, or locations, that store data or instructions. The amount of information physi-

cally stored in these locations varies from computer to computer. The term "word" usually defines the number of bits contained in each location, and is usually composed of 4, 8, 12, or 16 bits. A bit is a binary digit and is the smallest unit of data in a microcomputer. A word is the basic unit indicating the number of bits that are processed simultaneously. Each memory location has an address, and most computer memories can be accessed by means of these addresses in a random – rather than sequential – fashion. This is called RAM, for Random Access Memory. Read Only Memory, or ROM, can also be randomly accessed – but normally cannot be altered by the computer.

Unique addresses are also assigned to input/output devices enabling the CPU to access a specific I/O device.

GROUP 4. INPUT/OUTPUT (I/O). Computers must be interfaced in order to communicate with humans or with other devices. I/O devices are commonly called peripherals. Computer peripherals usually communicate in the same digital format as the computer. However, many situations exist in an analog – rather than digital – state, and must be translated into and out of a binary condition. This is usually accomplished through the use of Analog to Digital Converters (ADCs) during input and Digital to Analog Converters (DACs) during output.

Early input methods were IBM cards, paper tape, and magnetic tape. Today, the most popular human interface is the VDT, or Video Display Terminal, which uses a keyboard input and video screen output. Ultimately, computers will actually speak to and hear humans. Until then,

high-speed printers, disk and tape devices will remain the most numerous and widely-used output peripherals.

Modern computer terminals communicate information so rapidly that the process seems instantaneous to human perception.

All digital computers communicate and process messages by converting information to a system of binary numeric coding.

The previous course in Microcomputer Operations introduced the student to binary number systems and how their simple digital nature could be used to add and subtract numbers. This course, Microcomputer Hardware, will examine more closely electrical devices that enable a digital computer to add and subtract, as well as other more complex operations a digital computer performs. But first, a glimpse of the nature of electrical signal used by these devices is necessary.

DIGITAL SIGNALS

In a digital context, the term "signal" refers to symbolic meaning transmitted by means of periodic changes between two different electrical voltages against a pre-determined timeframe.

The notion of time in this definition points to another important element in the communication of a signal other than the voltage differential: the duration, or time span, of a signal.

Two terms that describe a signal's behavior in time are "level," and "pulse." A signal that remains constant at +5

volts through two consecutive time periods in order to communicate two consecutive digital states of 1 is called a level signal. This type of signal is normally used as a control signal that changes less frequently. A pulse signal communicates the same two consecutive 1 states as two rapid changes from 0 volts to 5 volts and back to 0. Pulse signals are normally used to indicate rapidly-changing data.

Very precise timing is required in digital computing. Durations in time normally will be measured at the precise moments when voltages change. The change from 0 volts to +5 volts (the change from the low state to the high state, also referred to as going high) is commonly called the leading edge of the signal. The return to 0 from +5 volts (the change from the high state to the low state, also referred to as going low) is commonly called the trailing edge of a signal. These points in time often trigger chain reactions between logic networks.

Within a computer, various channels are designated to communicate specific types of information. One such control channel, or control bus, sends a regular cycle of signals throughout the computer for timing purposes, providing a time-frame of relativity against which the duration of a signal may be measured.

In Figure 1, three channels containing digital signals are represented as they might be seen on the face of an oscilloscope through a fraction of a second. Notice that for each channel, two voltages are possible. (Although 0 and 5 volts are used, in practice a technician will encounter many voltages that may be used to communicate the binary states of 0 or 1, on or off, high or low, true or false).

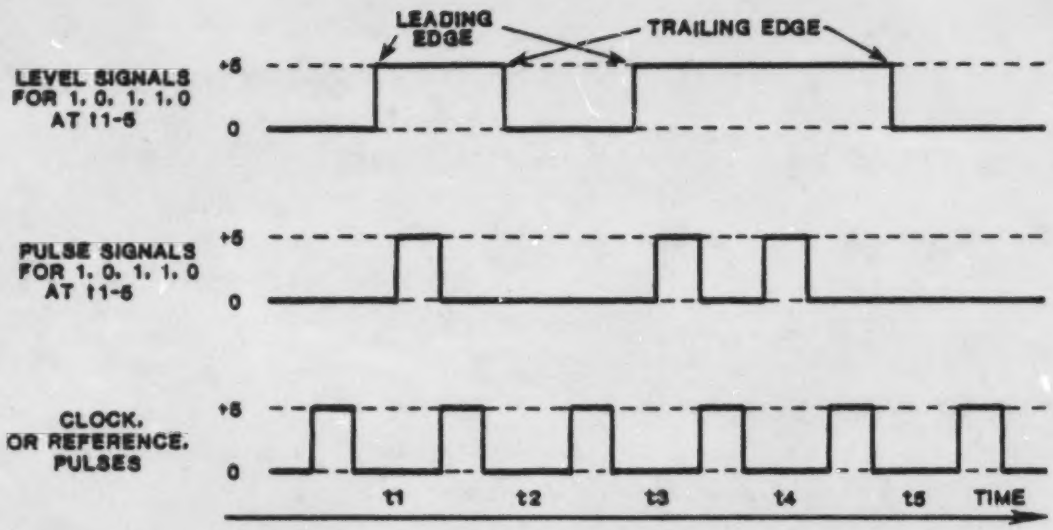


Figure 1. Two Channels Communicating Digital Information By Level or Pulse Signals, Against a Reference or Clock Signal.

It is helpful to imagine each clock period defining a window during which information is communicated, either by a pulse or by maintaining a level voltage during the clock period.

Clock control signals can occur on single or multiple buses. Multiple timing information may be generated from a single clock source and then communicated over a single line by a device called a phase splitter.

In Figure 2, a clock signal is represented as line T. Notice that changes in voltage occur over a short (non-immediate) period of time, resulting in a more diagonal waveform, as opposed to the square waveform. Figure 2, shows the clock line as two phased clock lines, A and B.

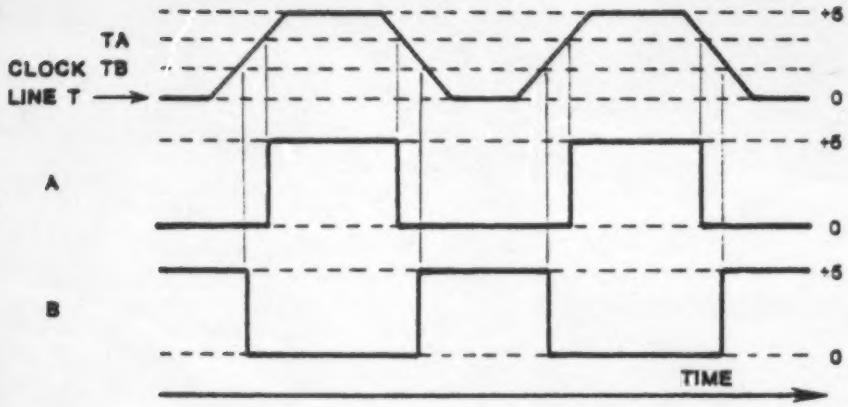


Figure 2. Two Phased Clock Lines Generated By a Single Clock Signal.

By utilizing devices sensitive to specific voltage levels, called threshold voltages, the signals on lines A and B may be produced. Signals on line A are high, or 1, when the clock voltage level is above the level indicated by line TA. Signals on line A are low, or 0, when the clock voltage level is below the level indicated by line TA.

Line B, however, is high when the clock voltage level is below the threshold level indicated by the line TB, and low when the clock voltage is above the level indicated by the line TB.

Notice that the two signals produced are mutually exclusive, in that signals are never high at the same time. In fact, there is a brief period of time when neither line is high. This delay can be useful in screening undesirable variations in the propagation characteristics of clock signals, called clock skew.

The term clock "cycle" refers to a defined repetition rate of continuous, consistent clock signals and is a impor-

tant measure of the internal speed of a computer. Frequency defines the number of cycles per second and is measured in Hertz.

Other types of buses used by the computer are the address bus, normally a unidirectional channel upon which signals flow in one direction; and the data bus, normally a bidirectional channel carrying information to and from the control unit.

Numbers may be transmitted within the computer, either sequentially or in parallel. For example, to communicate a byte (8 bits) of information sequentially, a clock cycle is divided into 8 clock periods per cycle. Each bit is transmitted in order, beginning with the Least Significant Bit (LSB) and ending with the Most Significant Bit (MSB). To communicate the same byte of information in parallel, eight separate circuits are required. Although this is more expensive in terms of circuits, the same byte can be transmitted in one clock pulse, or eight times faster than the sequential method.

Next, information on the specific logic networks used to communicate information will be introduced. The functions of these networks have not changed significantly since the earliest computers, though the electronic components used to construct these networks have changed dramatically. This module will deal with these logic networks and their function within a digital computer system. The next module will deal with the electronic components within each network.

Logic used in this module is commonly called positive true logic, because the most positive voltages in the examples cited represent the logic value 1, or true. Logic 0 corresponds to the 0, negative, or least positive voltage.

There is also a system of negative true logic which assigns logic 1 to the least positive, 0, or negative voltage. This system has advantages with some devices (from an engineering viewpoint, in order to optimize electrical performance), but such information lies outside the scope of this module.

GATES

Digital signals are generated and used by logic circuits – which are actually electrical switches, diodes, resistors, or transistors – arranged in a configuration that allows them to represent logical conditions.

Specific electronic components for each logic network will not be detailed because there are many possible combinations and components that will achieve outputs with the desired logic values. Each gate will have its own logic symbol representing discrete components, which will indicate function only.

These devices are called gates, due to their controlling function. Gates allow only certain types of signals to pass through the device while blocking others.

Initially, we will look at five simple – but important – logic functions: AND, OR, NOT, NAND, and NOR. Each will be discussed separately.

The AND Gate

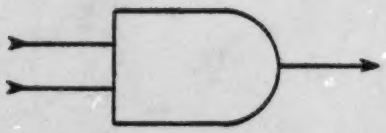


Figure 3. Logic Symbol for the AND Gate.

Figure 3 shows the symbolic representation of a 2-input AND gate. When both inputs entering the AND gate are 1, the output is 1. (For the remainder of this module, the terms 0 and 1 will be used to indicate binary states, and should not be confused with actual voltages). If either of the inputs is 0, the output from the device will be 0.

In Figure 4, the AND gate is represented as a series of two switches. If either switch is open, electricity will not pass through the series of switches; therefore, the AND gate's output will be 0. If both switches are closed, electricity will pass through the series and the AND gate output will be 1.

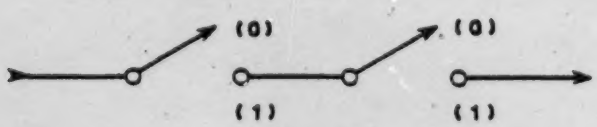


Figure 4. Series of Switches Representing an AND Gate.

A multiple-input AND gate functions in a similar fashion: all inputs must represent 1 for the output of the system to represent 1. Although the first sample logic gate is simple, it is helpful at this point to demonstrate the construction of a truth table (Figure 5) to summarize the consequences of

INPUT A	INPUT B	OUTPUT
0	0	0
1	0	0
0	1	0
1	1	1

Figure 5. Truth Table Depicting All Possible Input/Output Combinations of a 2-Input AND Gate.

all possible input conditions. Although there are only four possible conditions in a 2-input AND gate, these tables can be helpful in more complex situations.

If any inputs to an OR gate (Figure 6) are 1, the output is 1.

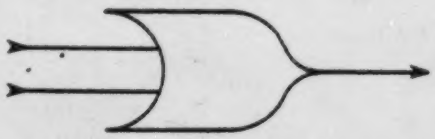


Figure 6. Logic Symbol for the OR Gate.

This can be represented by a series of switches in parallel, so that if any switch is closed (representing 1), electricity will flow through the system and its output will be 1. This is shown in Figure 7.

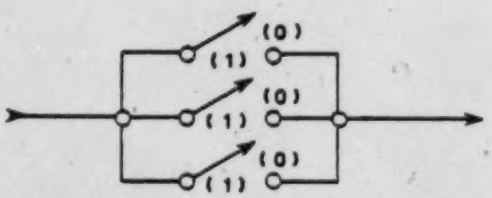


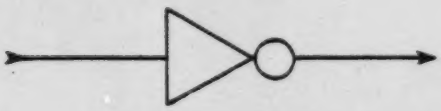
Figure 7. 3-Input OR Gate Depicted As 3 Switches Connected in Parallel.

Figure 8 shows the truth table for a 2-input OR gate.

INPUT A	INPUT B	OUTPUT
0	0	0
1	0	1
0	1	1
1	1	1

Figure 8. Truth Table
for a 2-Input OR Gate.

The logic symbol and truth table for a digital inverter (or NOT gate) are shown in Figure 9.



A	X
1	0
0	1

Figure 9. Logic Symbol and Truth Table
for a Digital Inverter.

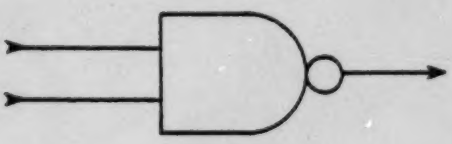
The small circle at the right point of the triangle in the logic symbol in Figure 9 is sometimes called an inversion bubble, and is the part of the logic symbol indicating the inversion function. Without the inversion bubble, the triangle would symbolize the identify function, representing circuitry that amplifies, isolates, and restores signal quality. However, with the inversion bubble, output from an inverter has the opposite, or inverse, value of the input.

The inversion bubble will also appear in the logic symbols of the next two devices, the NAND and NOR gates.

The NAND Gate

If any input to a NAND gate is 0, the output is 1. If all the inputs to a NAND gate are 1, then the output is 0. It is sometimes helpful to think of the logical function of this device by the words NOT AND, in other words, the inverse or opposite logic state of the output of the AND gate. Notice the similarity in the logic symbol for the AND and NAND gates, with the addition of the inversion bubble to the AND symbol indicating the inverting function.

The logic symbol and truth table for a two-input NAND gate are pictured in Figure 10. Notice that the result is the inverse, or opposite, of the truth table for the AND gate.

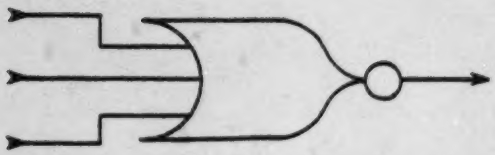


INPUT A	INPUT B	OUTPUT
0	0	1
1	0	1
0	1	1
1	1	0

Figure 10. Logic Symbol and Truth Table for a 2-Input NAND Gate.

The NOR Gate

If the inputs to a NOR gate are 0, the output is 1. If any of the inputs are 1, the output is 0. The relationship of the OR and NOR gates is similar to that of the AND and NAND gates, in that the outputs are exactly opposite as noted in the truth tables. The logic symbol and truth table for a 3-input NOR gate are pictured in Figure 11.



INPUT A	INPUT B	INPUT C	OUTPUT
0	0	0	1
1	0	0	0
0	1	0	0
0	0	1	0
1	1	0	0
1	0	1	0
0	1	1	0
1	1	1	0

Figure 11. Logic Symbol and Truth Table for a 3-Input NOR Gate.

From the above discussion, the following conclusions may be drawn:

1. The only condition that will allow the output of an AND gate, regardless of the number of inputs, to be 1, is if the value of all inputs is 1. In all other conditions, its output is 0.
2. The only condition that will allow the output of an OR gate, regardless of the number of inputs, to be 0, is if the value of all inputs is 0. In all other conditions, its output is 1.
3. The only condition that will allow the output of a NAND gate, regardless of the number of inputs, to be 0, is if the value of all inputs is 1. In all other conditions, its output is 1.
4. The only condition that will allow the output of a NOR gate, regardless of the number of inputs, to be 1, is if the value of all inputs is 0. In all other conditions, its output is 0.

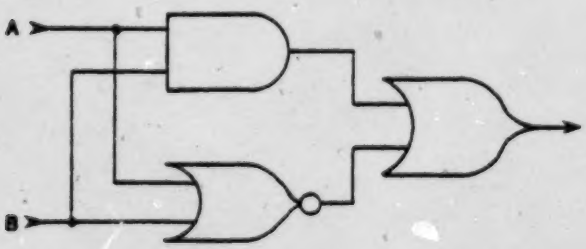
These conclusions summarize data presented to this point, and reflect logic conditions wherein only one precise condition changed the output from what it would be in any other case.

Outputs from the next logic devices to be presented respond in a different manner, indicating one state when their inputs are the same, and indicating the other state when their inputs are different.

EXCLUSIVE AND NON-EXCLUSIVE OR GATES

Some light sources in residences are controlled by two wall switches that function in the following manner: if both wall switches are up, or if both switches are down, the light is on; if either switch is up, and the other down, the light is off. This circuit functions like a NON-EXCLUSIVE OR gate, and is also called the exclusive NOR gate.

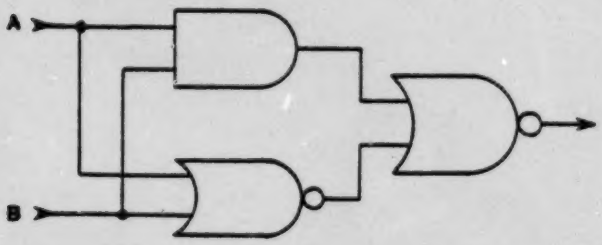
This circuit can be fashioned by utilizing an AND, OR, and NOR gate connected as shown in Figure 12, beside the truth table for a NON-EXCLUSIVE OR gate.



INPUT	INPUT	OUTPUT
0	0	1
1	0	0
0	1	0
1	1	1

Figure 12. Diagram of Logic Network Composing a NON-EXCLUSIVE OR Gate, with its Corresponding Truth Table.

An EXCLUSIVE OR gate functions exactly the opposite. If both light switches in the residence were up, or if they were both down, the light would be off; if either switch were up, and the other were down, the light would be on. The EXCLUSIVE OR can be made from an AND gate and two NOR gates, as depicted in Figure 13 beside its truth table.



INPUT	INPUT	OUTPUT
0	0	0
1	0	1
0	1	1
1	1	0

Figure 13. Logic Network and Truth Table for an EXCLUSIVE OR Gate.

To summarize, a NON-EXCLUSIVE OR gate indicates a 1 at output when its inputs are the same, a 0 when they are different. An EXCLUSIVE OR gate indicates 0 at output when its inputs are the same, a 1 when they are different.

The logic symbols for these devices are shown in Figure 14. Do not confuse the logic symbol for a NON-EXCLUSIVE OR gate with the logic symbol for a NOR gate.

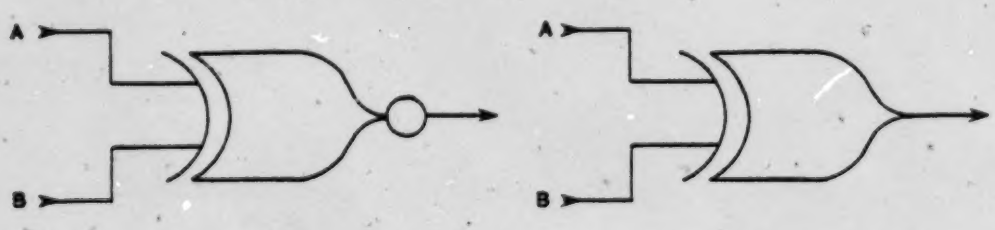


Figure 14. Logic Symbols for EXCLUSIVE and NON-EXCLUSIVE OR Gates.

Logic networks utilizing these basic gates may be combined in many different ways to produce results that match a desired truth table. Often there is more than one solution to the design aspects of a logic network, but most often the design using the fewest number of gates is the most practical.

ADDITION/SUBTRACTION CIRCUITRY

By combining the five simple logic circuits, the AND, OR, NOT, NAND, and NOR gates, the logic network now exists for the next – and very important – step in the capability of a digital computer. This step is the computer's ability to add, subtract, multiply, and divide numbers.

The main functional element of a digital computer is called an adder, and it comes in two basic types: the parallel adder, which will be introduced first; and the serial adder, which will be discussed after our introduction to register circuitry.

At this point, the student should recall from the first course in Microcomputer Operations the discussion on the binary number system and the addition and subtraction of binary numbers.

Subtraction, multiplication, and division are all modifications of binary arithmetic. In two's complement arithmetic, subtraction is achieved by complementing, or inverting (changing 0s to 1s and 1s to 0s), the number and adding binary 1. Multiplication and division are accomplished by shifting a number left or right, similar to the way in which the multiplication or division of a base 10 number by 10 is accomplished by moving the decimal point to the left or right.

Parallel adders are much faster than serial adders, providing an answer at almost the same instant that input signals arrive. But a very short length of time is required for the carry signal to ripple down through the adder circuitry. As a result, the parallel adder is sometimes called the ripple carry adder.

If any two numbers, expressed in their binary (hex) forms, were to be added together, one would first look at adding the two least significant bits (LSB) of each number, to which there would be four possible conditions, as indicated in Figure 15.

```
  0          1          0          1
 +0         +0         +1         +1
  0 (carry 0)   1 (carry 0)   1 (carry 0)   0 (carry 1)
```

Figure 15. Four Possible Conditions of Adding Two LSBs.

If a truth table were to be constructed for just this much of our addition problem, it would contain two inputs (one for each LSB) and two outputs (sum plus carry information). The truth table in Figure 16 below reveals that a carry condition exists only when both inputs are 1.

INPUT	INPUT	CARRY OUTPUT	SUM OUTPUT
0	0	0	0
1	0	0	1
0	1	0	1
1	1	1	0

Figure 16. Truth Table Indicating Four Possible Conditions of Adding Two LSBs.

The truth table for the carry portion of this addition is identical to the truth table for a 2-input AND gate. The truth table for the sum portion of this addition is identical to the truth table for a 2-input EXCLUSIVE OR gate. Not surprisingly, it is exactly these two gates that can comprise the half adder.

Actually, there are several combinations of logic networks that will result in the above logic results. Figure 17 shows a logic diagram for a half adder utilizing AND and NOR gates.

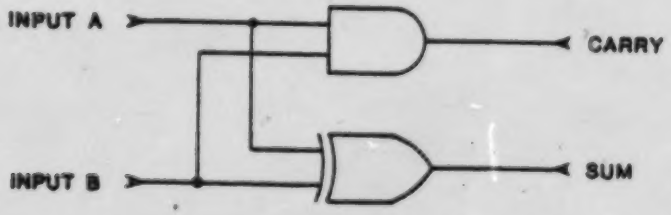


Figure 17. Logic Diagram of Half Adder, Utilizing AND and NOR Gates.

When the next two bits to the left in the binary number being added are summed, the carry information from the addition of the LSBs must be taken into account. Thus, the next addition will have 3 inputs, but once again, only 2 possible outputs. A 3-input, 2-output logic network may be formed from two half-adders and an OR gate, and is termed a "full-adder."

Rather than showing the full adder as a network of logic gates, the block diagram in Figure 18 summarizes the principles being utilized in adder networks. This same logic can be extended to form adder circuitry with the ability to add multiple-bit numbers together.

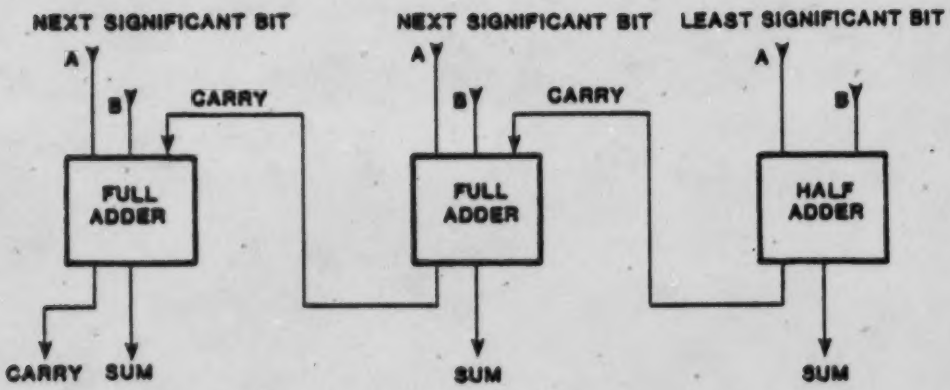


Figure 18. Block Diagram Showing How Adder Networks are Constructed.

DECODERS

Decoders use combinational logic to select a memory address or an I/O device. The decoding logic network enables a multiple number of addresses to be accessed by a few lines. For example, two lines will access four addresses, because there are four possible binary combinations that can be communicated by two lines. By using only 4 AND gates and 2 INVERTERS, it is possible to devise a network that will access four locations through two input lines. Three lines will access eight locations, and so on.

MULTIPLEXERS

Basically, a multiplexer enables many outputs to be connected to only one input, again through a network of the

basic logic gates, by selecting which output will pass through the network and into the input.

FLIP-FLOPS

Logic gates and adder circuitry are the simplest devices utilized by a digital computer to process signals by means of logic networks. The devices described to this point have been dependant on one or several inputs registering a binary state by means of a voltage (or absence of a voltage) that are meaningful only while these input voltage conditions are present. This is called combinational logic.

The concept of time, as introduced in the discussion of digital signals, requires sequential logic, implying the concept of storage of a signal through time.

Another type of logic device, called a flip-flop, enables the computer to store a signal for an extended time. These devices are the basic building blocks for memory registers, shift registers, and counters.

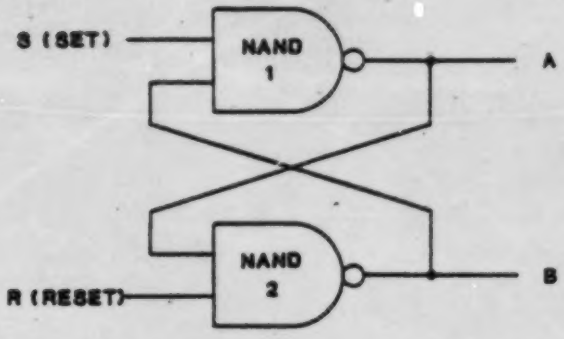
Although the ability to hold a condition through time in effect slows down the operation of a computer, this characteristic is necessary for the synchronization of signals. As signals are routed and altered within the computer system, they must sometimes be held in a particular location within the system until other signals necessary for certain operations arrive.

The most important characteristic of a flip-flop is that it tends to remain stable in one of only two conditions at any given time. For this reason, a flip-flop is sometimes called a bistable multivibrator.

The Set-Reset Flip-Flop

A simple set-reset flip-flop may be constructed by cross-coupling one of the inputs of two NAND gates to the outputs of each other, as shown in Figure 19, along with the truth table for this device.

Notice that the inputs are labeled "s" for set and "r" for reset, while the outputs are labeled "a" and "b." The logic values 1 and 0 are normally present at the outputs "a" and "b," and their positions are normally reversed (flip-flopped) by input pulses.



S	R	1	0
0	0	NO CHANGE	
1	0	1	0
0	1	0	1
1	1	NOT USED	

Figure 19. Diagram and Truth Table for a Crossed NOR Flip-Flop.

The following is a detailed description of the logic involved in a flip-flop circuit, and the student is urged to have a clear understanding of the concepts involved in this device before proceeding to new material, as many of the devices studied in this course revolve around this basic circuit.

Assume for the moment that the logic value 1 is present at the inputs labeled "s" and "r" in Figure 19. If either output "a" or "b" is also at logic value 1, then the opposite

output must register logic 0. In actuality, "b" will always be the inverse of "a," and "a" will always be the opposite logic value of "b."

For example, if output "a" was 1, and input "s" was 1, then the other input to NAND gate 1 must be a 0; otherwise, output "a" could not logic 1. (If needed, consult the truth table requirements for a 2-input NAND gate.) As shown in Figure 18, the value on this input to NAND gate 1 is identical to output "b"; therefore, "b" must also have logic value 0. Furthermore, if "a" is 1, then both inputs to NAND gate 2 are 1 (remember "r" is assumed 1 for the moment) and again output "b" is 0, as was predicted.

On the other hand, if output "b" was 1, and input "r" is 1, then the other input to NAND gate 2 must be 0. If it is 0, then it follows that "a" is also 0.

By convention, the state of the flip-flop is indicated by the logic value at output "a." If "a" is 1, the flip-flop is said to be "set" or at logic 1; if "a" is 0, the flip-flop is said to be "reset" or at logic 0. Labeling the inputs "s" and "r" for a set-reset flip-flop is fairly standard. Students will encounter different variables used to indicate the outputs, such as X or Q for "a," but since the value at the opposite output is the inverse, opposite, and complement of the value at "a," the "NOT bar" from Boolean algebra will be drawn over the variable used for output, "b," as in $\bar{X}$ or $\bar{A}$.

Going back to the above description, assume that both "s" and "r" are still at logic 1, while "a" is 1 and "b" is 0 (the "set" state). If a 0 pulse occurs momentarily at "r," then the inputs to NAND gate 2 become 0 and 1, and the output "b" goes from 0 to 1. When this occurs, output "a" goes to 0, for the reasons outlined above. Therefore, a 0 pulse mo-

mentarily at "r" has reset the device. Notice that if the momentary 0 pulse had occurred at "s" instead of "r," nothing would have happened. This would have been a signal to the device to go the set state, but it already was in the set state, and would have remained so.

Therefore, the following conclusions can be drawn concerning a flip-flop. The flip-flop has only 2 possible stable states: when one of the outputs is 0, the other is 1, and vice-versa. Once the device has been set, it will remain in that condition until it receives a signal to go to the reset state. Once reset, the device will remain so until it receives a signal to go to the set state. A flip-flop is also called a latch.

The memory properties of a flip-flop circuit are apparent from the above discussion, but the device is not without its faults. If 0 were to occur on both "s" and "r" at the same time, the state reflected by the flip-flop should remain unchanged. If 1 occurs on both inputs "s" and "r" at the same time, the result would be indeterminable. Since the goal of the design of the flip-flop is to be able to store two and only two determinable states from one or more inputs, the following discussion will show how the addition of extra gates or multiple flip-flops will form logic networks that have the desired characteristics.

This is an important point, because desired characteristics change from computer design to computer design and application to application. Certain logic devices are more suitable in certain situations. However, both the function of logic devices and the hardware that composes them are adaptable to the specifications of the designer.

The Trigger Flip-Flop

In the set-reset device described above, a set impulse (once the flip-flop has already been set) will result in no change. Likewise, a reset pulse, once the flip-flop has been reset, will result in no change. The addition of a third input, labeled "t" in the block symbol for a trigger flip-flop in Figure 20, will add a new capability. A trigger pulse (in other words, a pulse on the input labeled "t") will change a flip-flop to the other state, regardless of the state it is in at the time the pulse is received. Thus, the logical consequence of a trigger pulse is to invert the value stored in the flip-flop.

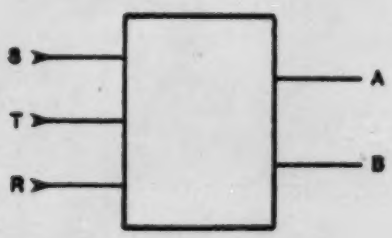


Figure 20. Block Symbol for a Trigger Flip-Flop.

A trigger flip-flop can exist without a set, reset capability. This device is also called a toggle, and simply alternates the logic values 1 and 0 at the outputs "a" and "b."

The Master-Slave Flip-Flop

To avoid the use of capacitors, which are not easily adaptable to modern miniaturized electronic components, a

trigger flip-flop is placed ahead of a crossed-NOR flip-flop to store (through a short period of time) the input voltages that are applied to the crossed-NOR (or set-reset flip-flop) circuit.

These additions help to preserve the critical timing tolerances, coordinated by the master clock pulses, that enable a computer to perform its tasks with a higher degree of precision.

Computers are equipped with safety devices to check serial parity, a method of detecting the loss or addition of a digital pulse in a series of bits that make up a word.

Briefly, this can be accomplished by constructing a simple logic device containing trigger flip-flops and AND gates. If each one-bit pulse on the data line toggles a trigger flip-flop on and off, and the total number of bits in the word is an even number, the number of toggle pulses should be even. Therefore, the flip-flop will be in the reset state at the end of a word. If the number of toggles pulses is odd, the flip-flop will be in the 1 state thereby indicating serial parity check.

The Steerable Flip-Flop (J-K)

In some logic situations, a flip-flop circuit might be logically dependent on itself, requiring the computer to sense the status of the flip-flop circuit prior to altering its contents. In such situations, oscillations might occur that cause the flip-flop to change states constantly during a single clock period. It would not be possible to predict which state the flip-flop would be in at the end of that period.

The flip-flop circuits described so far have the ability to respond to set and reset pulses, and to invert upon receiving a trigger pulse. These devices do not have the ability to indicate 1 at the "a" output by changing the state of the flip-flop, if need be, or by leaving it alone if it already indicates the desired output. This capability can result by utilizing a 3-input AND gate to control the master flip-flop. The inputs to an AND gate utilized for this special purpose are customarily labeled "J" and "K" (the third input is a clock pulse). These circuits are sometimes called J-K flip-flops.

D-Type Flip-Flops

A D-type flip-flop has the most simple relationship between the state it will assume and its input. When enabled by the clock line, a D-type flip-flop will take on the value appearing upon its input line.

Register

As mentioned earlier, microprocessors process digital words (varying in length from machine to machine, but not within a microprocessor) that are processed in parallel, rather than in series.

There are many types of register networks. Some have only information storage capability. Others, like shift registers, not only store information but pass information to the right or left. Figure 21 pictures a simple shift register

that shows the effect input has on the various output lines as it is shifted through time.

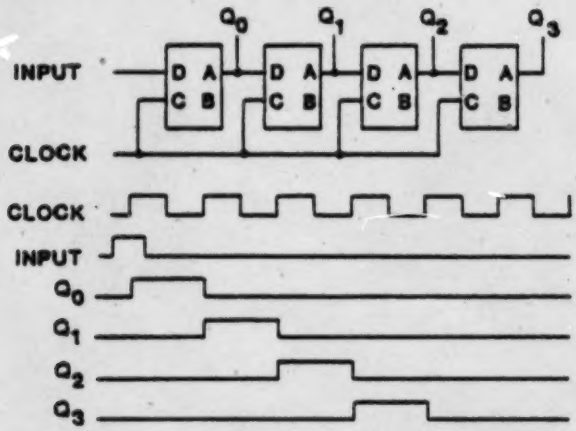


Figure 21. A Simple Shift Register and its Outputs Through Time.

Counters

Another type of register is called the counting register or simply counter. J-K flip-flops are easily adapted to divide-by-two use, in which the output frequency is exactly half that of the input frequency to the network. Figure 22 illustrates a network of three J-K flip-flops that form a simple three-stage counter. This diagram clearly illustrates the increment function of a counter, as each time period counts up in binary code. Counters may also be constructed to decrement, or count down.

Notice that each rate of change is half the rate of the previous time period (denoted $T_1 - T_3$ above). The table sum-

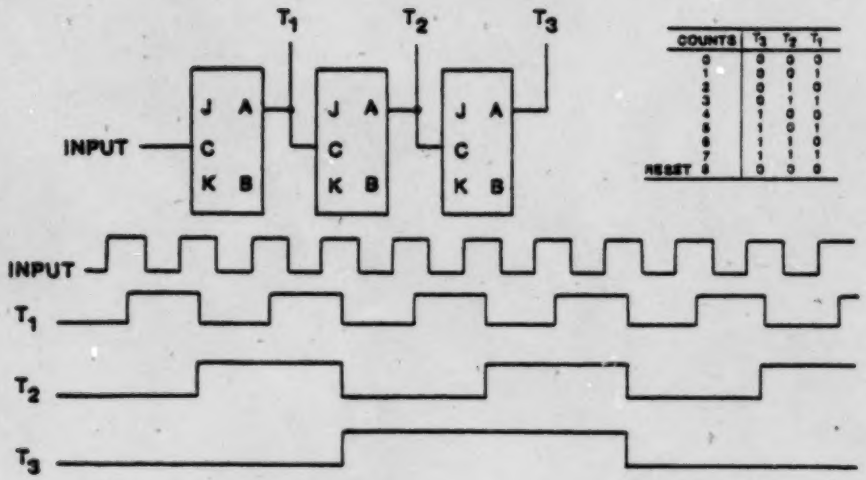


Figure 22. Divide-By-Two Counter Using Three Connected J-K Flip-Flops.

marizes the outputs on $T_1 - T_3$ at each count period. Notice that the result is the binary equivalents of the decimal numbers 1-7.

Buffers

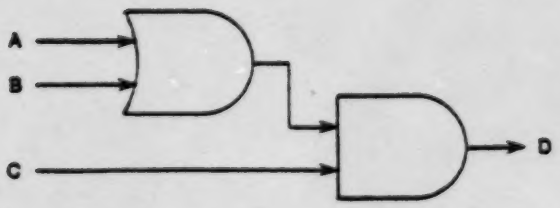
The term "buffer" indicates a specific type of memory device that functions as a temporary holding tank for information. It is required because of the substantial difference in the rate at which an input device feeds information that is to be processed into the computer compared to the high speed at which actual processing occurs. A buffer is also used to receive and hold data not complete enough for the central processing unit to utilize.

EXERCISES

1. Above the clock pulse line depicted below, draw a waveform that illustrates:
   a. The binary number 100110 as a series of level signals.
   b. The binary number 010111 as a series of pulse signals.

+5
0
+5
0
+5
0

2. Develop a truth table for the logic network shown below:



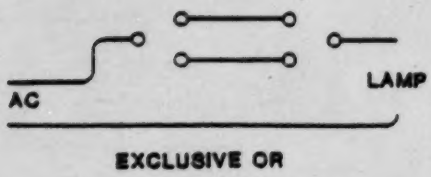
A *	B *	C *	D
0	0	0	
1	0	0	
0	1	0	
0	0	1	
1	1	0	
0	1	1	
1	0	1	
1	1	1	

3. Develop a logic network for the truth table shown below:

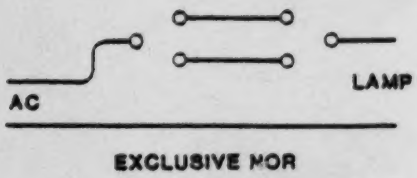
A *	B *	C *	D
0	0	0	0
1	0	0	0
0	1	0	0
0	0	1	1
1	1	0	1
0	1	1	1
1	0	1	1
1	1	1	1

(Draw logic network here.)

4. In the module, the functions of the AND and OR gates were explained by using a simple diagram of switches in an electrical circuit, whose open and closed conditions matched those of the logic condition required for those gates. Using the beginning provided below, develop a diagram depicting the EXCLUSIVE OR and EXCLUSIVE NOR gates as switches in a circuit.

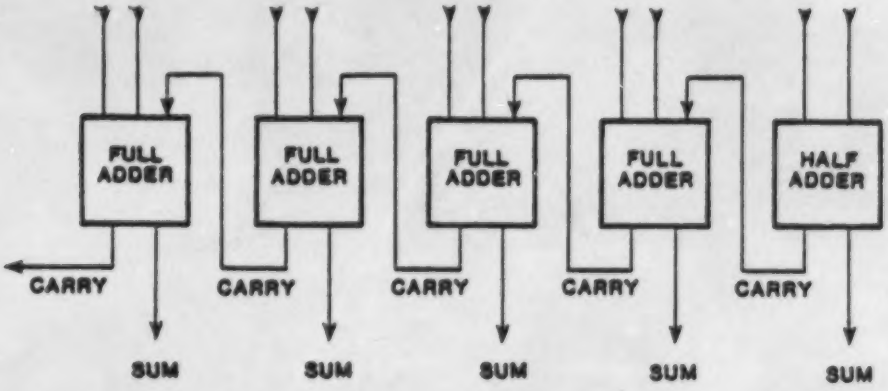


EXCLUSIVE OR



EXCLUSIVE NOR

5. Show the conditions of each input and output of a full adder circuit during the addition of the following binary numbers: 00101, 01011.



LABORATORY MATERIALS

Power: +5 volts at 1 amp and +12 volts at 0.1 amp
KIM-1 Microcomputer
ICs:
- 7400 Quad 2-input positive NAND gate.
- 7402 Quad 2-input positive NOR gate.
- 7404 Hex inverter.
- 7405 Hex inverter, open collector.
- 7474 Dual D-type positive edge-triggered flip-flop with preset and clear.
- 74161 Binary counter, synchronous, preset input (asynchronous clear).
- 74125 Quad gated buffer-three state.
- 7485 4-bit magnitude comparitor, separate A-B output.

LABORATORY PROCEDURES

PROCEDURE 1

First, examine the operation of these ICs by applying the specified voltages to the inputs of one of the gates and testing the output with a VOM. If needed, chart the outputs of the various input combinations to prove that they agree with the truth tables studied. The schematic and specifications for each chip will be listed on the package. Should any of the ICs not function normally, use them as questionable ICs to test on the KIM.

PROCEDURE 2

Refer to the manual on Microcomputer Operations, Module 1, for instructions to set up the KIM-1 microcomputer.

Next, utilize the IC test program provided on the cassette tape to check the ICs above.

To read a tape into the KIM:

1. Place 00 into address 00FL.
2. Place the ID number in address 17F9. The IC test program is ID #04.
3. Start executing at address 1873.
4. Advance the tape to the beginning of program 04. Temporarily disconnect the lead from the tape recorder to the KIM so that the tape will be audible. Listen for four distinct audio signals. These will indicate the beginning of program 04. Re-connect the lead from the output of the tape recorder to the KIM. Turn the volume of the tape recorder to near

maximum. Check to make sure that you have 12 volts connected to the KIM.

5. Start the tape reading. If four 0s appear, the tape has been successfully read. If four Fs appear, or if nothing happens after a few minutes, the tape has not been properly read.

This program can be used to compare a questionable IC with one known to be functioning. To do this, the computer's outputs are connected to the inputs of each IC, and IC's outputs connected to the computer's inputs. When a good IC is used, the microcomputer learns how the IC operates by running through all possible logical input values and storing the resulting outputs from the IC. Then, if the good IC is replaced by a questionable one and the stored values are compared to the output of the questionable IC, any discrepancy indicates a faulty circuit.

Locate the IC output pins from the schematic on the package and connect them in any order to lines in the computer port B (PB). Locate the IC inputs and connect them to lines in port A (PA). Start execution at address 0200. The display should be all zeros.

Press PC to have the KIM learn what outputs a good IC should generate. Now substitue a questionable IC, making sure that the outputs and inputs from the second IC correspond to those of the first, and press DA. A display of all zeros indicates a good IC.

If the IC is bad, the left pair of digits give the output on PB, the center pair give the actual output and the right pair give the output recorded for the good IC. You must also consider the possibility that the first IC whose operations was taken to be normal could possibly have been faulty.

REFERENCES

Abrams, Marshall D. and Stein, Phillip G. Computer Hardware and Software. Reading, MA: Addison-Wesley, 1973.

Barna, Arpad and Porat, Dan I. Introduction to Microcomputers and Microprocessor. New York: John Wiley and Sons, 1976.

Baron, Robert C. and Piccirilli, Albert T. Digital Logic and Computer Operations. New York: McGraw-Hill, 1967.

Bartee, Thomas C. Digital Computer Fundamentals. New York: McGraw-Hill, 1971.

Bell, Leonard. Digital Concepts. Beaverton, Oregon: Tektronix, Inc., 1968.

Flores, Ivan. Computer Logic. Englewood Cliffs, NJ: Prentice-Hall Inc., 1960.

Givone, Donald D. and Roesser, Robert P. Microprocessor/ Microcomputers: An Introduction. New York: McGraw-Hill, 1980.

Kershaw, John D. Digital Electronics: Logic and Systems. North Scituate, MA: Duxberry Press, 1976.

1. Define the following terms:
   a. Leading edge.

   b. Flip-flop.

   c. Register.

   d. Buffer.

   e. Program counter.

   f. Primary memory.

2. Identify the following abbreviations:
   a. VDT
   b. TTL
   c. LSI
   d. CPU
3. Draw the logic symbols for the following:
   a. The AND gate.

b. The NOR gate.

c. The EXCLUSIVE-OR gate.

d. The NOT gate (INVERT).

4. Explain the difference between the INCLUSIVE, the EXCLUSIVE and the NON-EXCLUSIVE OR gates.

5. True or False:

_____a. The only condition that will allow the output of an AND gate, regardless of the number of inputs, to be 0, is if the value of any of the inputs is 1.

_____b. The table depicting the 4 possible conditions resulting from the addition of two binary corresponds to the truth table of the EXCLUSIVE OR gate.

6. A simple set-reset flip-flop may be constructed by cross-coupling the inputs of 2: (circle correct answer)
   a. AND
   b. OR
   c. NOR
   d. NAND gates.

7. In negative true logic, which of the following voltages in each pair represents the true condition: (circle correct answer)

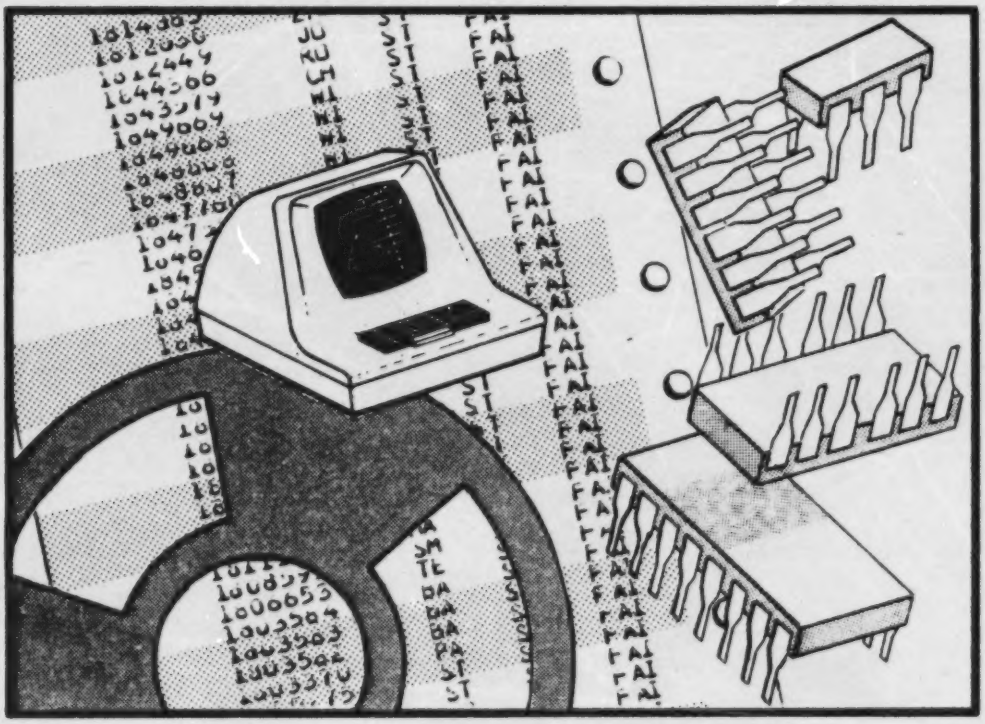
a1.	+5	b1.	-2	c1.	0
a2.	0	b2.	-5	c2.	-5

8. Why is parallel processing more desirable than serial processing?

ENERGY TECHNOLOGY

CONSERVATION AND USE

MICROCOMPUTER HARDWARE



MODULE MH-02

SEMI-CONDUCTOR
LOGIC FAMILIES

TECHNICAL EDUCATION RESEARCH CENTER - SOUTHWEST
4800 LAKEWOOD DRIVE, SUITE 5
WACO, TEXAS 76710

INTRODUCTION

Many different components are combined with a computer to achieve digital logic. Since the 1970s, these components have been almost exclusively solid-state semiconductor devices, normally found in largely-integrated configurations. This module reviews the characteristics of the most common solid-state logic families, and provides a brief overview of semiconductor electronics.

PREREQUISITES

The student should have completed Module MH-01, "Digital Components."

OBJECTIVES

Upon completion of this module, the student should be able to:

1. List the advantages and disadvantages associated with the selection of a specific semiconductor logic family with regard to a defined application.
2. Recognize and identify the circuit symbols for solid-state devices.
3. Define the following terms:
   a. N-type silicon.
   b. P-type silicon.
   c. PN junction.
   d. Diode transistor logic (DTL).
   e. Transistor.
   f. Bipolar.

g. Semiconductor.
h. Emitter.
i. Base.
j. Collector.
k. PNP transistor.
l. NPN transistor.
m. Saturation.
n. Cutoff.
o. Conductivity.
p. Breakdown voltage.
q. Leakage current.
r. Maximum power dissipation.
s. Sink load.
t. Source load.
u. Unit load.
v. Pull-up resistor.
w. Fan-out.
x. Fan-in.
y. Schottky TTL.
z. Emitter-coupled logic (ECL).
aa. Integrated injection logic (IIL or I^2L).
bb. Merged transistor logic (MTL).
cc. Metal oxide semiconductor (MOS).
dd. Field-effect transistor (FET).
ee. PMOS.
ff. NMOS.
gg. Complementary metal oxide semiconductor (CMOS).
hh. FET source.
ii. FET drain.
jj. FET gate.

ATOMIC MAKEUP OF SILICON CRYSTALS

The solid-state revolution that occurred in electronics during the late 1960s – and especially the 1970s – was based on silicon crystals and their use in integrated circuits.

Because the world of semiconductor IC devices is small, this discussion begins at the atomic level, since basic electricity is no more than the movement of electrons.

The nucleus of a silicon atom contains 14 protons, whose positive charge is negated, or offset, by 14 negatively-charged electrons orbiting the nucleus at speeds approaching the speed of light.

Protons and electrons have opposite charges of equal strength. When an atom contains an equal number of protons and electrons, charges cancel each other out, and the atom is said to be electrically neutral.

In any atom, the electrons that orbit the nucleus tend to group at specific distances from the nucleus. These orbits are called shells, and each shell has a tendency to contain a specific number of electrons. When the outer orbit of an atom is filled with this specific number of electrons, the atom assumes a very stable, balanced condition. At this point, it becomes difficult to dislodge an electron from the atom. However, if a single electron winds up in the outer shell of an atom that prefers eight electrons in its outer shell, this lone electron can be easily moved from atom to atom. Since like charges repel, and opposite charges attract, an electron flow can be set up between

atoms having lone electrons. This flow is electricity. If a moving electron encounters an atom with seven electrons in an outer shell that requires eight, the electron is captured and becomes part of a stable atom.

Materials composed of atoms with lone electrons make excellent electrical conductors. Stable atoms whose outer shells are full make excellent insulators against the flow of electricity.

Of the 14 electrons, the 10 electrons in the lower orbits of a silicon atom are stable. The remaining four electrons, located in the outer orbit, reside in a shell that requires eight electrons to be stable.

If a group of silicon atoms is formed in such a way that each atom shares one of its four outer electrons with its neighbor (in return for one of its neighbor's outer electrons), the resulting matrix is a stable condition. By sharing electrons (Figure 1), the silicon atom has fulfilled its requirement for eight electrons in its outer shell. Silicon atoms joined by bound, or shared, electrons have crystalized. This silicon crystal is not a good conductor of electricity. A representation of silicon crystal is shown in Figure 1 below. (The inner circle in each atom in Figure 1 represents the nucleus and 10 lower orbit electrons.)

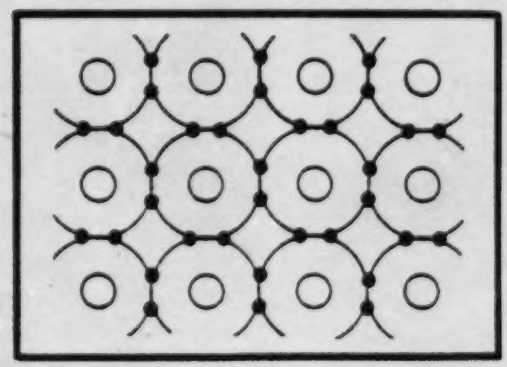


Figure 1. Shared Outer Electrons of Silicon Atom for a Crystal Lattice.

If other materials are introduced into pure silicon when it is still molten, a substance may be formed which has the desired electrical property.

Phosphorous atoms have 15 protons in each nucleus, and five electrons in their outer shell. If phosphorous atoms are introduced into molten silicon, they combine with silicon atoms to form silicon crystal, leaving one extra, lone electron looking for a place to go. Since this new material has extra negative charges, it is termed "N-type" silicon. The silicon crystal is now able to conduct electricity.

Boron atoms have only three electrons in each outer shell, and when combined with silicon, produce crystals with a spot for one missing electron. These holes are very important in solid-state electronics, as they are generally thought to flow in the opposite direction of electrons. Since this new boron/silicon material has a deficiency of electrons, it is positively charged, and is termed "P-type" silicon.

DIODES

A diode is a device that allows electricity to flow in one direction while blocking the flow of electricity in the opposite direction. A solid-state semiconductor diode may be constructed by placing a tiny chip of P-type silicon against a small chip of N-type silicon. The point at which these two regions meet is called the PN junction. In Figure 2, a PN semiconductor diode is diagrammed beside its electrical symbol. Note that the arrow in the electrical symbol for a diode points in the direction <u>opposite</u> the

electron flow. In all diodes, the negative lead is also called the cathode, while the positive lead is also called the anode.

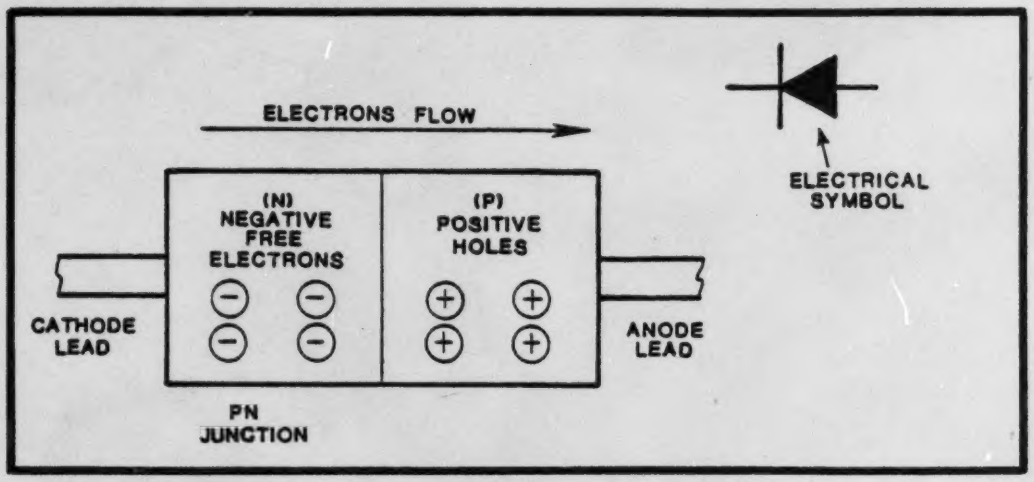


Figure 2. Diagram and Electrical Symbol for a PN Diode.

If an electric current enters the N region of the diode, the negatively-charged electrons which make up that current begin to repel electrons in the N region toward the PN junction. Since it is assumed that this is a closed circuit, current also drains from the P region of electrons, thereby creating new holes. This forces existing holes toward the PN junction, where they are filled by the abundant supply of electrons now flowing toward the PN junction within the N region. In other words, when current is flowing from cathode to anode, the PN diode allows current to flow.

(Note: The term "current" in this module will refer to "electron" current and not "conventional" current, which flows from positive to negative, unless notated as conventional current.)

When current is applied in the opposite direction, negative electrons are drawn away from the PN junction toward the cathode in the N region. On the other side, electrons are being sent into the P region, where they are plugging holes. This action attracts holes away from the PN junction toward the anode.

Thus, the area close to the PN junction becomes like pure silicon crystal which, as was stated before, is not a good conductor of electricity. In this manner, a PN diode blocks the flow of electricity in one direction, while allowing electricity to flow in the opposite direction.

Figure 3 below illustrates how PN diodes may be used to construct the AND and OR gates described in Module MH-01. PN diodes are required to isolate situations where one input is high and the other is low. Without PN diodes, electricity would flow in one input and out the other – in the OR gate, for example.

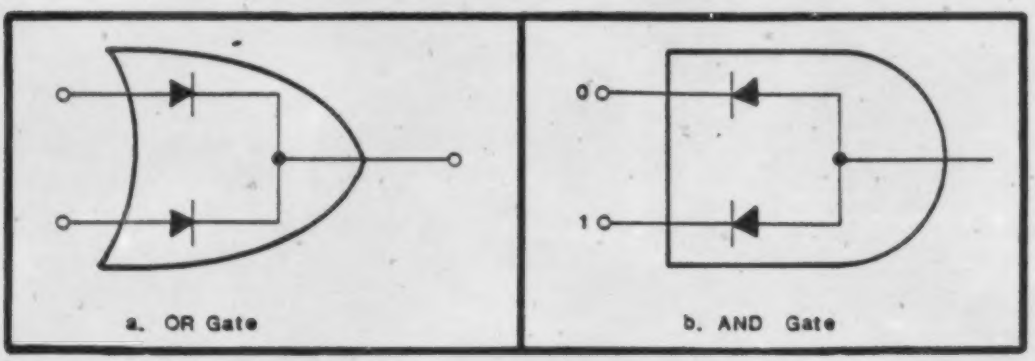


Figure 3. Logic Gates Formed by Diodes.

Constructing logic gates with diodes is rather crude. Diodes consume power without restoring any signal. This effect multiplies when many gates are interconnected, resulting in deteriorating signal quality. The output from a

diode logic gate could be restored by attaching it to a transistor, which can also serve as an inverter to turn the diode AND gate into a NAND gate with better output quality.

When diodes and transistors are used together to form logic networks, the process is commonly called diode transistor logic (DTL). However, transistors can perform the same logic functions as diode-based circuits, and transistor-transistor logic (TTL) is the most widely-used form of semiconductor bipolar device. The term "bipolar" refers to semiconductors whose primary current flows through regions made up of both positively and negatively-polarized materials. The term "semiconductor" originates from the fact that these devices sometimes conduct – and sometimes block – the flow of electricity.

TRANSISTORS

A transistor is an active electrical device which utilizes a control current to affect the condition of a control area, called the base, that in turn governs the flow of a larger current between the emitter and the collector.

N and P-type silicon may be put together in two configurations to accomplish this transistor action.

An NPN transistor utilizies two N regions (one a collector and the other an emitter) separated by a narrower P region, which serves as the base. Figure 4 depicts a NPN transistor.

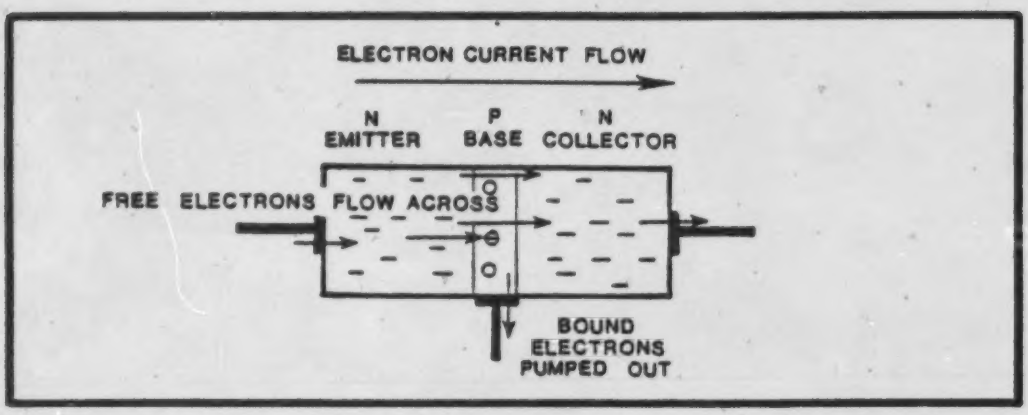


Figure 4. An NPN Semiconductor Transistor.

If a current is passed from the emitter to the collector, flowing through the base, electrons will begin to plug holes in the P-type material (provided the control circuit is open so that no electrons are being drawn out of the P region). This gradually builds a negative charge in the P region, which eventually strengthens to the point that electrons trying to pass through the P region are repelled. This negative charging of the P region happens rapidly, causing the transistor to shut off any current flow from emitter to collector in about 50 nanoseconds.

If a current is now passed through the base from the emitter in the control circuit, electrons are forced out of the P region in the base, slightly lowering the barrier of negatively-charged electrons in the base that prevent current flow. Normally, from 50 to 100 electrons in emitter/collector current are allowed to flow for every electron removed from the base by the control circuit. In this fashion, transistors use small currents to control larger ones.

PNP transistors follow the same principle of NPN transistors, but the base material of a PNP transistor is comprised

of N-type silicon, while the emitter and collector are comprised of P-type silicon. Electricity flows in the opposite direction in a PNP transistor, from collector to emitter, as indicated in the electrical symbols for PNP and NPN transistors in Figure 5 below.

The arrows in the transistor symbols indicate the emitter lead, and point in the opposite direction of electron flow. The vertical lines respresent the base, and the lead opposite the emitter is, of course, the collector.

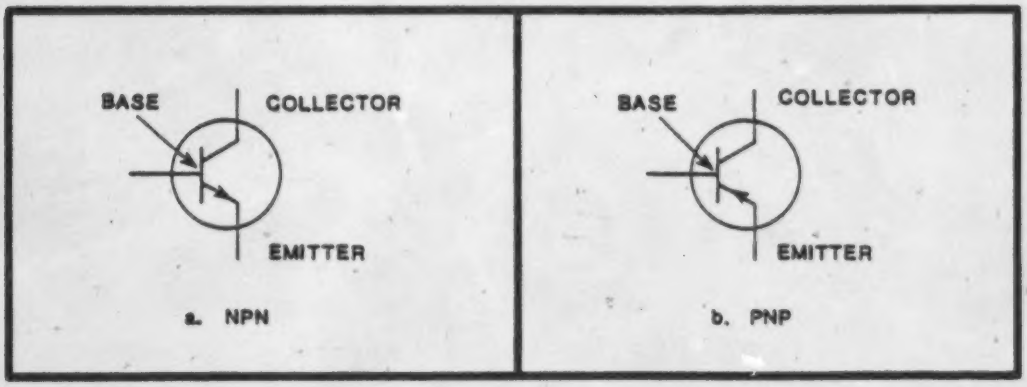


Figure 5. NPN and PNP Transistor Circuit Symbols.

The condition within a transistor that results from the build-up charge in the base area preventing emitter/collector current flow is called cutoff. There is also a condition in transistors that occurs when the control element reaches a certain level. The transistor is in effect wide open, and emitter/collector current flow is limited only by external circuit factors and not by the transistor. This condition is called saturation and occurs in silicon transistors at around 0.6-0.7 volts.

The speed at which a transistor fluctuates between saturation and cutoff constitutes its operating speed when

used as a switch. This rate is one of the important specifications used to select a particular type of transistor for a circuit. Transistors switch at much greater speeds than devices composed of diodes, relays, or vacuum tubes. Figure 6 illustrates the DTL NAND gate referred to earlier, and also depicts a TTL version of a NAND gate.

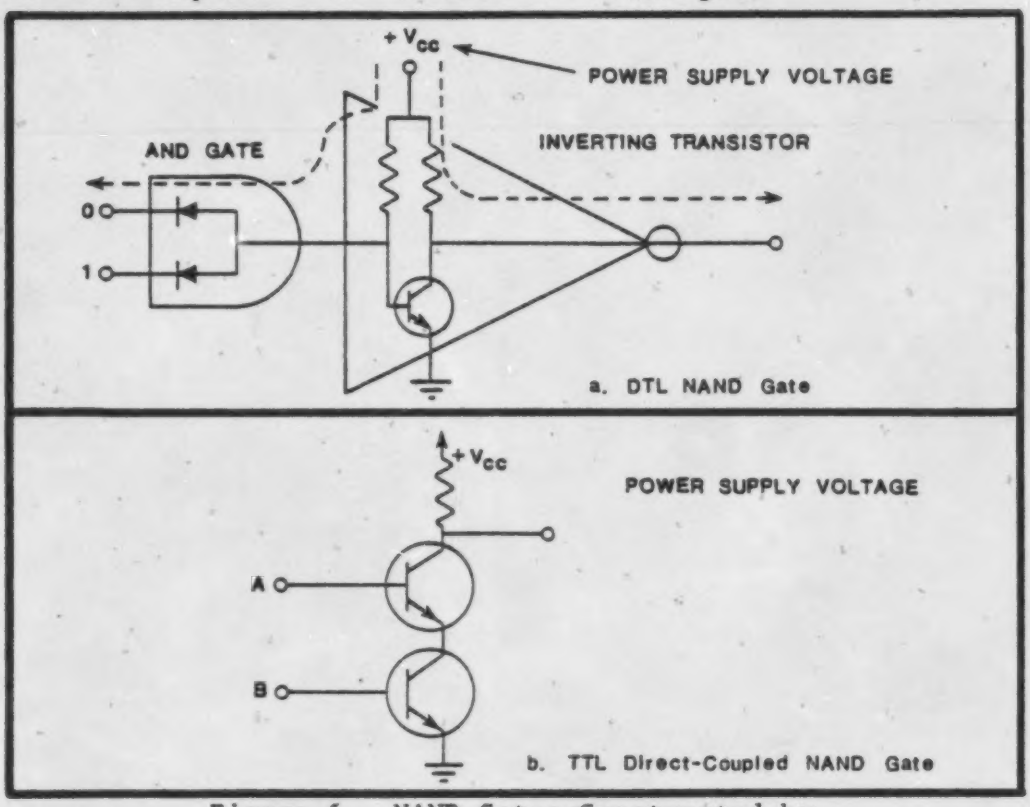


Figure 6. NAND Gates Constructed by Diode Transistor Logic (DTL) and Transistor-Transistor Logic (TTL).

Some other specifications used to rate transistor performance are as follows:

1. Conductivity – a measurement of how easily a current moves through a transistor when it is ON.

2. Breakdown voltage – the highest voltage level that a transistor can resist in its OFF state.
3. Leakage current – current that flows when it should not, usually a small amount along a circuit path that should be blocked.
4. Maximum power dissipation – power that is transformed to heat and radiated away, and is therefore wasted energy.

It is important for the student to understand that logic networks involving transistors are not perfectly precise circuits in terms of exact logic 1 and 0 levels. The demands of mass producing tiny IC semiconductor devices result in slight variations in the performance characteristics of the circuitry, but these are held within the defined specifications of the manufacturer.

There are many factors that contribute to a certain amount of imprecision in semiconductor logic circuitry, such as noise, variations in rise or fall times, and propagation delay time.

The sharp, well-defined edges of the digital signals (termed "leading edge" and "trailing edge") in Module MH-01 are described as having important timing functions within the computer.

Figure 7 illustrates this ideal signal against a more accurate representation of a signal encountered in an actual computer environment.

Noise, or minor fluctuations in both voltage and current levels, is the result of thousands of simultaneous electrical events occurring close to a specific signal.

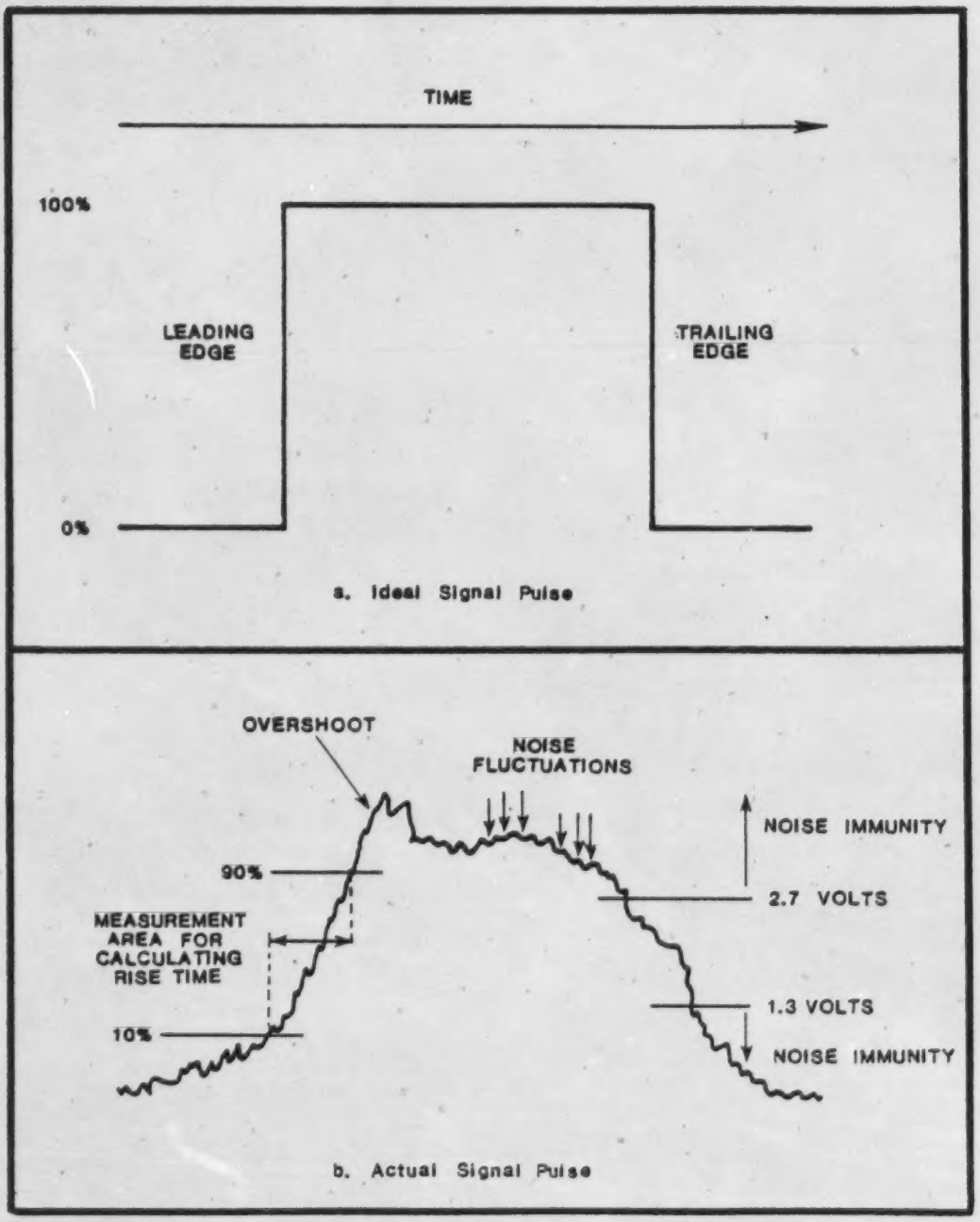


Figure 7. Comparison of Real and Ideal Digital Signals.

The voltage levels at the right of the actual signal in Figure 7b define a narrower window than the signal voltages used in most transistor logic networks. This is one way of providing some protection against noise spikes, or rapid fluctuations in voltage due to noise, that occur when the signal is at logic 1 or at logic 0.

Since it takes a certain amount of time for a transistor to reach saturation level or fall to cutoff level, this actual signal might represent the output signal of a transistor, showing the rise and fall times. These times are sometimes measured between 10% and 90% of the waveform differential, because of the possibility of overshoot or other imprecisions. This measurement is shown on the left of the actual signal in Figure 7b.

Before the characteristics of each family of semiconductor logic devices are detailed, several other items encountered in conjunction with transistor networks will be defined briefly.

There are two types of current loads associated with the voltage levels in logic circuits. The term "sink" load may be defined as the positive current (usually measured in mA) that an input will draw when the output of the gate driving it is at a specified logic 1 level. Sink load currents typically lower the voltage level of the driving output when it is logic 1, but its effect on the driving output during logic 0 is negligible. The term "source" load current adds energy to the system it is driving, having little effect on logic 1 inputs and tending to raise the voltage level of a logic 0. TTL devices normally use source loads.

The load current most representative of the majority input circuits in a logic system is called the unit load.

The unit load is used to compute easily the number of logic devices that may be driven from the outputs of a specified logic device within the computer system.

TTL LOGIC DEVICES

As mentioned before, transistor-transistor logic devices are the most widely used bipolar semiconductor electronic components. Many newer technologies that have been developed remain TTL-compatible, as the specifications for TTL circuitry are uniform throughout the industry.

Logic values of 1 and 0 are represented in TTL devices by voltages between +2.4 volts and +5 volts for logic 1, and voltages less than +0.4 volts for logic 0. These are consistent for both standard and low-power TTL devices.

The current (normally in mA) input level in standard TTL for logic 0 is normally 1.6 mA, while the output levels are normally 16 mA or 10 times greater. This means that standard TTL devices have a fan-out of 10, or in other words, one output of 16 mA will drive 10 inputs at 1.6 mA.

Low-power TTL devices have a fan-out of more than 20, with normal inputs for logic 0 at 0.36 mA and outputs at 8 mA. Because low-power TTL devices use approximately one-fourth the current of standard TTL, they are extremely useful in situations that demand low current drain. Low-power TTL devices have one-tenth the maximum power dissipation, meaning that much less power is wasted as heat, but they are also about one-third as fast as standard TTL.

If outputs from several TTL logic gates are connected to pull-up resistors (or resistors that ensure that a logic level remains high unless forced to a logic 0 state) and

outputs connected together, a new logic gate has, in effect, been superimposed over the existing two gates without any additional circuitry. This is called wired logic, and is illustrated in Figure 8 as two "ANDed" NAND gates.

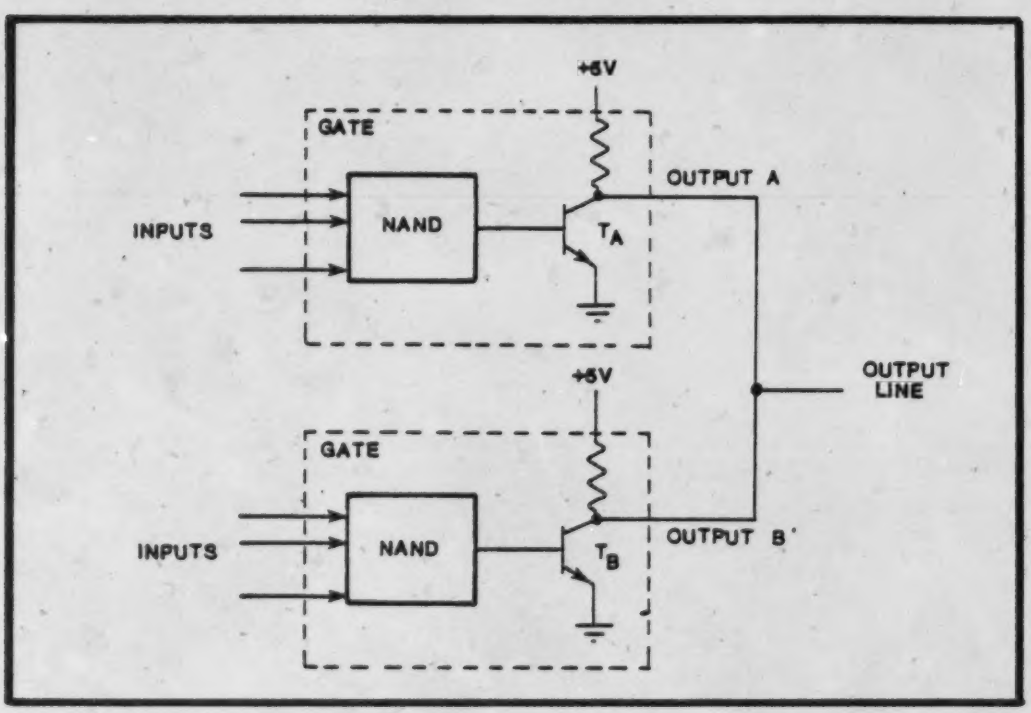


Figure 8. Circuit Diagram for a Wired Logic Gate Performing the AND Function on Two NAND Gates.

Several advantages result from wired logic. The additional logic function is essentially free and, due to its construction, does not lessen the speed of the device much.

Since there is no limit on the number of gates that may be connected this way, a fan-in technique may be used to create bus lines that can communicate signals between devices by either unidirectional or bidirectional means.

In order to use wired logic with pull-up transistors

at the outputs of the devices to increase switching speed, a third state – different from either logic 1 or 0 – would be required to prevent damage to the internal transistors of gates registering opposite outputs.

TTL circuits with three-state outputs register a high impedance or high Z state during the absence of a logic 1 or a logic 0-drive output.

A variation of the semiconductor diode previously described, called a Schottky diode, is formed by substituting a metal for the P-type silicon in a PN diode.

Electron orbits reflect the energy level of electrons; higher energy states result in electrons being located in more outer orbits.

There is a substantial differential between the energy levels of the outer electrons in N-type silicon and the metal used in Schottky diodes. This differential can be exploited in a Schottky device, producing an improved diode capable of higher switching speeds at lower voltage levels.

When a Schottky diode is used with a standard transistor, as illustrated in Figure 9, the Schottky diode acts to lower the level at which a transistor reaches saturation from around 0.6-0.7 volts to 0.2-0.3 volts. This greatly increases the speed at which the transistor can alter its states between saturation and cutoff, the ON and OFF.

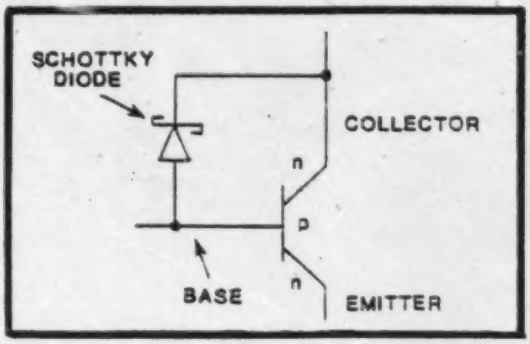


Figure 9. Schottky Diode Across a Collector Base Junction in an NPN Transistor.

Notice in Figure 9 the distinctive electrical symbol for the Schottky diode.

Schottky TTL also has a standard and low-power form, but because some of the circuitry utilizes Schottky diodes, the circuitry resembles DTL rather than TTL devices. But circuits are TTL-compatible, and the important factor is that operating speeds are increased three to four times.

EMITTER-COUPLED LOGIC (ECL)

Another bipolar logic family is called emitter-coupled logic, or ECL. ECL transistors avoid saturation, and thereby increase switching speed, by carefully controlling the collector current.

ECL technology is newer than TTL, and is not as widely used. ECL circuits are not TTL-compatible, but are the fastest switching circuits yet developed. Of semiconductor technologies, ECL has the best noise immunity but also has the highest power dissipation.

INTEGRATED INJECTION LOGIC (IIL or I^2L)

The newest, and most promising, semiconductor technology is called integrated injection logic, or I^2L. I^2L devices take advantage of multiple transistors sharing the same P-type silicon for their collectors and bases. Innovative techniques such as these are sometimes called merged transistor logic (MTL) and these techniques hold forth the promise of the best of each of the prior-developed semiconductor technologies.

MOS TECHNOLOGIES

All semiconductor devices and families discussed so far have been bipolar in nature, utilizing two distinct types of oppositely-charged silicon.

The remaining semiconductor logic families, though not the newest, were the first devices to achieve large scale integration (LSI) to create a true computer-on-a-chip. These families are built around a metal oxide semiconductor, or MOS.

MOS technology is made possible through the use of field-effect transistors (FETs). A field-effect transistor is similar to transistors already described, in that it utilizes a control current to vary resistance between two terminals.

Two regions of P-type silicon, called the source and the drain, are set into either side of an N-type silicon substrate, called the channel. A metal electrode, called a gate, lies above the entire length between the source and the drain. The gate is separated from the source and drain by a thin insulator.

The MOSFET operates much like a PNP transistor, with the source functioning as an emitter, the gate a base, and the drain a collector.

However, the controlling action of the transistor is accomplished through voltage – rather than current – by utilizing the effect of the electrical field formed in the gate.

When voltage is applied, electrons in the N-type channel are forced away from the gate. The remaining holes provide a bridge for current from the source to the drain.

A simple representation of this transistor is shown in Figure 10.

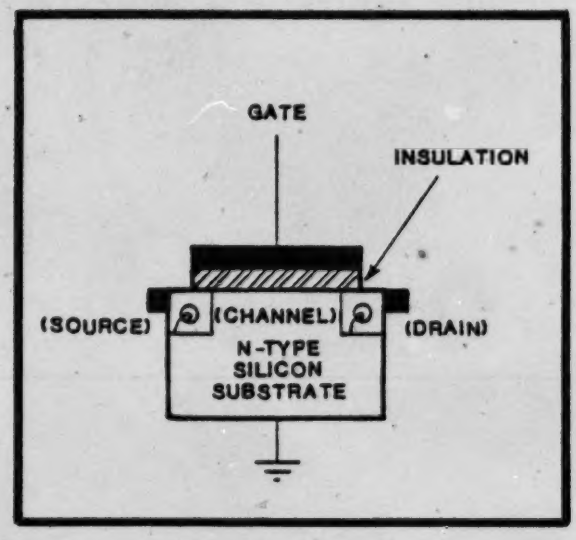


Figure 10. An NMOS Field-Effect Transistor.

MOSFETs that use N-type silicon as the channel are termed "NMOS." A similar device using P-type silicon as the channel with N-type silicon for the source and drain is termed "PMOS." Devices utilizing both transistor types in the same circuit are called complementary metal oxide semiconductors, or CMOS.

PMOS

PMOS was the original metal oxide semiconductor family. PMOS devices can be formed with very high densities, but unfortunately, PMOS devices operate at low speeds. As a result, PMOS devices have been used widely in calculators, as well as microprocessors.

Some PMOS devices are TTL-compatible, but typically require only one-fourth of the physical area of TTL devices. However, TTL devices are, on the average, 10 times faster than PMOS devices.

NMOS

NMOS technology is newer than PMOS. Even greater advances were made in LSI with these devices.

While more expensive than PMOS, NMOS devices are twice as fast and almost one-third as large.

Most NMOS devices are TTL-compatible, and there are many widely-used NMOS microprocessors available today from numerous manufacturers.

CMOS

CMOS devices are the most recent development in MOS technology. CMOS devices have achieved wide popularity in situations that utilize their ruggedness, high noise immunity, and low power requirements.

CMOS devices are much less dense than their NMOS or PMOS counterparts, taking up almost as much physical area as TTL.

TTL devices are still faster than CMOS, but CMOS is twice as fast as NMOS. CMOS is difficult to interface with TTL, due to variations in the types of inputs required.

EXERCISES

1. Identify the following abbreviations:
   a. TTL.
   b. DTL.
   c. MTL.
   d. LSI.
   e. MOS.
   f. ECL.
   g. IIL.
   h. CMOS.
2. Draw the electrical circuit diagrams for the following devices:
   a. PN junction diode.
   b. NPN transistor.
   c. Schottky diode.
   d. PNP transistor.
3. Fill in the blank.
   a. Electricity is the movement of _____.
   b. The terminals of a FET that correspond to the base, emitter and collector are _____, _____ and _____.
   c. Silicon atoms with shared electrons form a _____.
4. Rate the following devices 1, 2, 3 in terms of speed and density.

	Speed	Density
a. CMOS	_____	_____
b. PMOS	_____	_____
c. NMOS	_____	_____

5. The newest semiconductor technology is ...
   a. TTL.
   b. DTL.

c. I^2L.
d. ECL.

6. Name as many of the important criteria possible that are used to rate the performance of a transistor.
7. Draw a diagram to illustrate the function of a ...
   a. PN junction.
   b. Diode OR gate.

LABORATORY MATERIALS

Power supply: +5 volts.
IC: 7410 TTL integrated 3-input NAND gate.
Various unmarked NPN and PNP transistors (markings may be covered for lab).
VOM.

LABORATORY PROCEDURES

The following laboratories offer the student experience in breadboarding circuits that utilize transistors and checking circuit outputs against various input conditions. A procedure is also offered for discovering if an unmarked transistor is a NPN or PNP transistor.

LABORATORY 1. BREADBOARDING CIRCUITS.

Figure 11 shows the diagram for a 7410 3-input NAND. By following the steps below, two of the gates will be configured similarly to the wired logic similar to Figure 8. What is the difference?

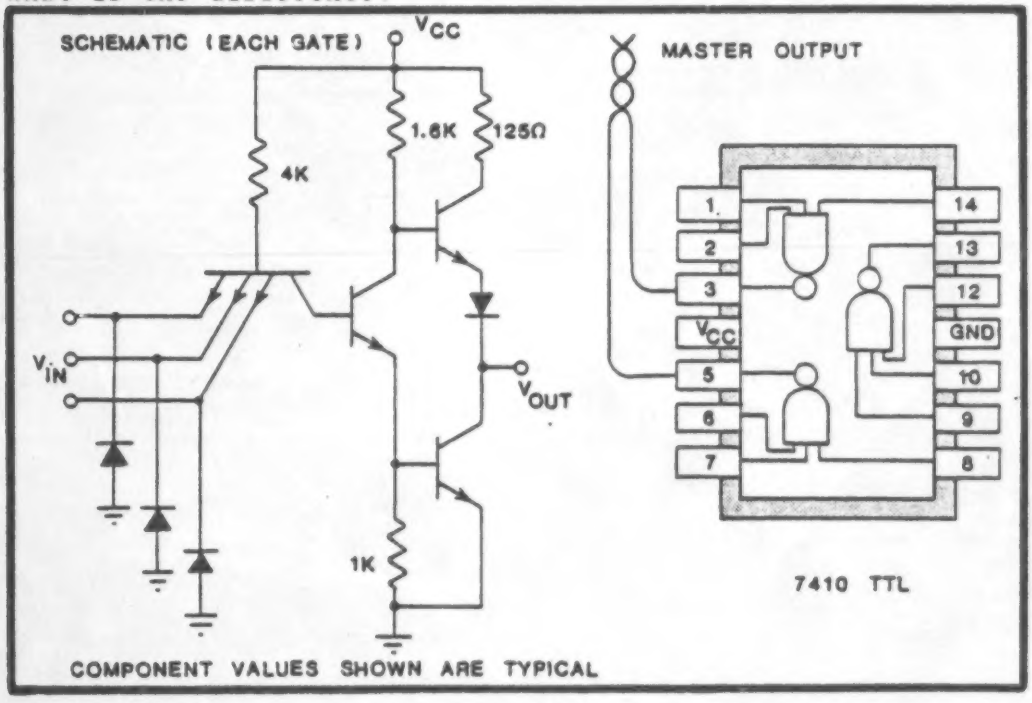


Figure 11. TTL Integrated Three-Input NAND Gate.

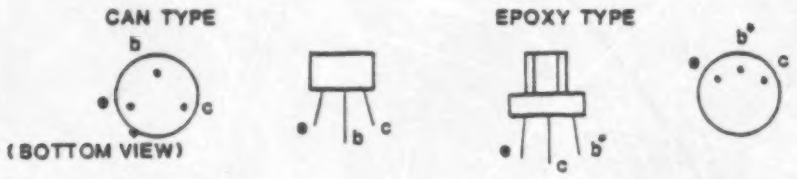
1. Connect the ground connection labeled GND to the ground on the power supply.
2. Attach wires to output pins 5 and 3, and connect the ends of the wires together. This point will be called master output.
3. Connect +5 volts to pin labeled V_{cc}.
4. Using the VOM, apply +5 volts to each of the inputs

according to the chart on the next page, recording the corresponding voltage output in the spaces provided in Data Table 1. In addition, record the values for current and resistance in Data Table 1.

LABORATORY 2. IDENTIFYING NPN AND PNP TRANSISTORS.

The following steps will allow the student to identify a NPN or PNP transistor.

1. Identify the emitter, base, and collector. Use the diagram below:



*Watch for this difference from some manufacturers.

2. Place the lead from the + terminal of the ohmmeter to the collector lead and the - terminal to the base lead. Record the resistance as R(a) in Data Table 2.
3. Reverse the leads and measure again. Record the resistance as R(b) in Data Table 2.
4. If R(a) is greater than R(b), it is a PNP transistor. If R(b) is greater than R(a), it is a NPN transistor. (If the resistance values are equal, or if they measure an infinite value, throw the transistor away!) Record the transistor type in Data Table 2. This procedure may be repeated several times.

DATA TABLE 1. BREADBOARDING CIRCUITS.

INPUT CONDITION:

Pin #:	1	2	14	6	7	8
1.	+5	+5	+5	+5	+5	+5
2.	+5	0	+5	+5	+5	+5
3.	+5	+5	+5	0	0	+5
4.	0	+5	0	+5	0	0

RESULT:

Pin 3:			Pin 5:		
Current	Voltage	Resistance	Current	Voltage	Resistance
_____	_____	_____	_____	_____	_____
_____	_____	_____	_____	_____	_____
_____	_____	_____	_____	_____	_____
_____	_____	_____	_____	_____	_____

Master Output:

Current	Voltage	Resistance
_____	_____	_____
_____	_____	_____
_____	_____	_____
_____	_____	_____

DATA TABLE 2. IDENTIFYING NPN AND PNP TRANSISTORS.

	R(a)	R(b)	Transistor Type
1.			
2.			
3.			

REFERENCES

Bartee, Thomas C. Digital Computer Fundamentals. New York: McGraw-Hill, 1977.

Gilmore, Charles M. Beginner's Guide to Microprocessors. Blue Ridge Summit, PA: Tab Books, 1977.

Kershaw, John D. Digital Electronics: Logic and Systems. North Scituate, MA: Duxbury Press, 1976.

Leventhal, Lance A. Introduction to Microprocessors: Software, Hardware, Programming. Englewood Cliffs, NJ: Prentice-Hall, 1978.

Soucek, Branko. Microprocessors and Microcomputers. New York: John Wiley and Sons, 1976.

Texas Instruments Learning Center. Understanding Solid-State Electronics. Dallas, TX: Texas Instruments, Inc., 1978.

TEST

1. Explain the functions of the emitter, base, and collector in a transistor.

2. Define the terms "saturation" and "cutoff" with respect to the operation of a transistor.

3. Why is it important that semiconductor technologies are TTL-compatible?

4. Explain what happens when electrons try to flow from the anode to the cathode in a PN junction diode.

5. Why is a transistor called an active device as opposed to a diode?

6. What is noise and where does it come from in an electrical signal?

7. What are the voltages that indicate logic 1 and 0 in TTL devices?

8. How does a Schottky diode improve the function of a transistor?

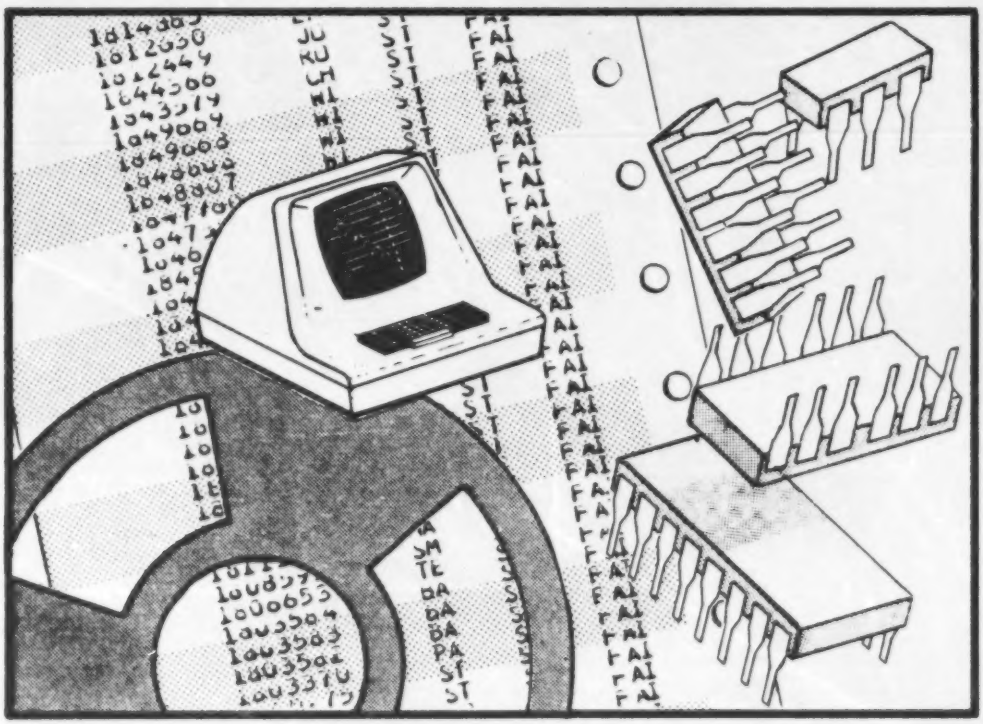
9. Explain why a field-effect transistor is well-named.

10. Why are the rise and fall times of a signal significant?

ENERGY TECHNOLOGY

CONSERVATION AND USE

MICROCOMPUTER HARDWARE



MODULE MH-03

INPUT/OUTPUT DEVICES
AND TECHNIQUES

TECHNICAL EDUCATION RESEARCH CENTER - SOUTHWEST
4800 LAKEWOOD DRIVE, SUITE 5
WACO, TEXAS 76710

INTRODUCTION

Most microcomputers utilize special circuitry and techniques to communicate information to and from the outside world. This module is a brief introduction to those circuits and techniques, as well as input and output peripherals that are used in the communication process.

PREREQUISITES

The student should have completed Modules MH-01, "Digital Components," and MH-02, "Semiconductor Logic Families."

OBJECTIVES

Upon completion of this module, the student should be able to:

1. Identify serial and parallel transmission techniques.
2. Define the difference between synchronous and asynchronous transmission or reception.
3. State the function of many of the available input/output (I/O) peripherals.
4. Define the following terms:
   a. I/O ports.
   b. Handshake.
   c. Universal asynchronous receiver transmitter.
   d. Universal synchronous receiver transmitter.
   e. Baud.
   f. Interrupt.
   g. Polled interrupt.

h. Vectored interrupt.
i. Flag.
j. Priority encoder.
k. Isolated I/O.
l. Memory mapped I/O.
m. Incorporated I/O.
n. Selector.
o. Decoder.
p. Multiplexor.
q. Keyboard or switch bounce.
r. Cathode-ray tube (CRT).
s. Video display terminal (VDT).
t. Optical character reader (OCR).
u. Impace printer.
v. Binary coded decimal (BCD).
w. Extended binary coded decimal interchange code (EBCDIC).
x. American standard code for information interchange (ASCII).
y. Light-emitting diode (LED).
z. Floppy disk.

INPUT/OUTPUT DEVICES AND TECHNIQUES

The first two modules dealt with the logical functions of a microcomputer and examined some of the hardware that are used to implement logic.

This module examines how the computer communicates with the outside world through input/output (I/O) processes.

Several reasons combine to make the study of input/output devices somewhat complicated.

First, there are almost as many different types of I/O devices as there are computer applications. I/O devices may be simple switches or single-channel devices that communicate only the current temperature of a room. I/O devices may be complex printers, cathode-ray-tube terminals – or more recently, digital voice synthesizers that speak.

Secondly, the speeds at which I/O devices operate vary widely. Some devices change one bit every three hours. Some devices change at the rate of hundreds of thousands of bits per second (bps).

Finally, the various signals that communicate with the outside world can differ in voltage, current, and resistance levels; number of channels required; method of change, and so forth.

Logically, the terms "input" and "output" are relative to the microcomputer. Input is information available to the microcomputer by means of an input port, or a physical connection to an outside information source. Likewise, output is information available from the microcomputer by means of an output port, or a physical connection to an outside receiver. Both ports are normally under the control of the

microcomputer, although the exact method of control can vary, as will be shown.

I/O devices (or peripherals) can be electromechanical, mechanical, or electronic. If their output signals are not in digital form, they must be converted by means of an analog-to-digital converter (ADC or a-d converter). If the peripheral will accept only analog input, the microcomputer must convert its digital output to analog by using a digital-to-analog converter (DAC or d-c converter). These devices will be covered in Module MH-04, "Analog/Digital Conversion." The hardware required to enable a computer and a peripheral to communicate is called an interface, and specific interfacing devices and methods are detailed in Module MH-06, "Bus Interfacing."

I/O ports may be assigned specific addresses, just like memory, so that the microprocessor can select one of a number of I/O ports. (The term microprocessor refers to the microcomputer central processing unit [CPU] rather than all the elements of a microcomputer.) Peripherals can even be used as memory storage devices. Many times, the only tangible difference between internal storage and peripheral storage to the microprocessor is the substantial difference in the time it takes to physically move the signal through available channels to the peripheral memory device.

PARALLEL/SERIAL TRANSMISSION

Parallel transmission enables the computer to communicate with a peripheral over multiple channels at the same time. Usually, the number of channels corresponds with the number of bits in the word length used by the microprocessor.

Obviously, this method of data transmission can communicate the same information considerably faster than can be accomplished by using a single line, which requires serial transmission.

In addition to the fact that the capacity to move signals between a computer and a peripheral is greater with parallel connections, serial data must be grouped into the proper word length by a converter device after it is received into the microcomputer. This additional operation is not required by the faster parallel transmission method.

Figure 1 shows the difference between the way an 8-bit word, or byte, is transmitted in serial and parallel form. Start and stop bits added to the serial signal are for synchronization purposes.

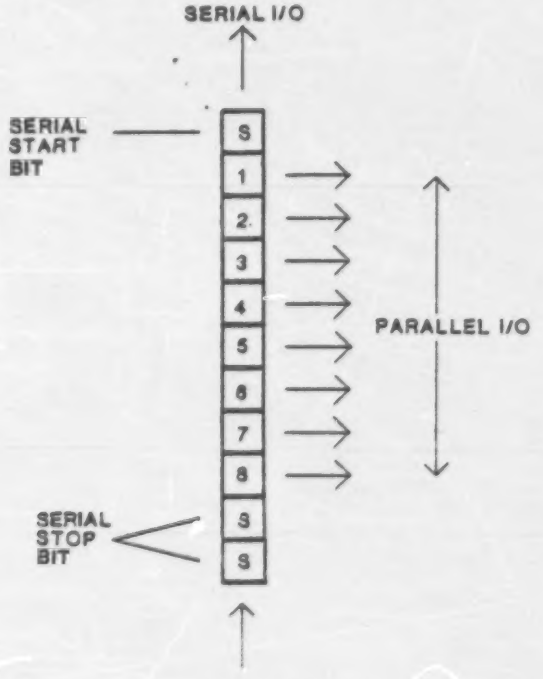


Figure 1. Two Methods for Sending an 8-Bit Word.

SYNCHRONOUS/ASYNCHRONOUS COMMUNICATION

One critical problem in I/O is peripheral/microcomputer data exchange.

The microcomputer can receive information from the peripheral at a rate that is different from the peripheral's rate of sending information.

Likewise, the rate of peripheral reception of data varies from the rate at which a microcomputer sends data.

Therefore, systems had to be developed that either stored or transformed the data – or somehow influenced the operating mode of the peripheral.

There are three basic methods for coordinating I/O. They are the following:

1. Regular sampling occurs when the design application of the system allows the computer to look for a change in I/O at a specified rate. This method is normally used with peripherals that are always ready, such as simple switches. Normally, data transfer rates are slow.
2. Asynchronous communication occurs when specific control signals, either on separate lines or as start and stop bits, provide synchronization for data transfer. This can occur at varying rates.
3. Synchronous communication occurs when an I/O signal is defined by its position in time (as defined by an external clock). Because these signals are continuous and require no other synchronizing signals (other than the clock signals), this method achieves the highest I/O transfer rates.

Regular Sampling

The front panel of a computer is normally a series of switches that allow addresses and data to be entered into the microcomputer in binary code. Lamps, or light-emitting diodes (LEDs), indicate the result, also in binary ON/OFF coding.

A problem occurs when a microprocessor attempts to read a logic level produced by a switch that is in the process of changing. Certain electronic trigger devices improve the rates at which switches change states. But in an application such as the front panel example, this problem is solved through special switches, labeled load or deposit. The computer samples these special switches at regular intervals, and when it finds them enabled, performs the data transfer operation indicated by the positions of the switches.

Obviously, the sample rate at which the computer looks for an enable condition should be greater than the minimum amount of time between changes in condition. In the case of human-operated switches, this rate is extremely slow compared to a microcomputer that can perform thousands of operations a second.

Another method of solving the problem of reading the output from a changing switch is through programming, or software. The computer's controlling software can provide a method for reading the output from a switch several times, then comparing readings for accuracy.

Asynchronous Communication

Even most peripherals that operate at faster-than-human speeds cannot provide data as fast as a microcomputer can read it. In most cases, the microcomputer will be waiting on the peripheral to send additional information while it performs other operations.

If temporary storage registers constructed from flip-flop circuits are placed at the I/O ports, the microcomputer can get the information to be input when it is ready, while the peripheral can get the output information when it is ready. An additional channel can communicate control status information, such as "ready to receive new information" and "new information ready to be sent." The resulting system enables data to be exchanged at varying rates between a microcomputer and a peripheral, as shown in Figure 2.

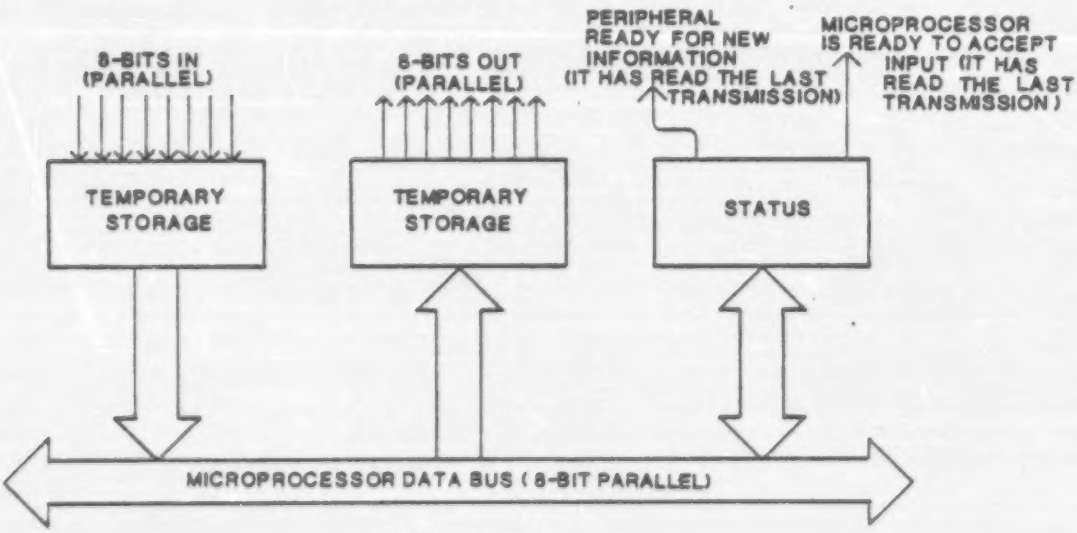

8-BITS IN (PARALLEL)
8-BITS OUT (PARALLEL)
PERIPHERAL READY FOR NEW INFORMATION (IT HAS READ THE LAST TRANSMISSION)
MICROPROCESSOR IS READY TO ACCEPT INPUT (IT HAS READ THE LAST TRANSMISSION)
TEMPORARY STORAGE
TEMPORARY STORAGE
STATUS
MICROPROCESSOR DATA BUS (8-BIT PARALLEL)

Figure 2. Parallel I/O Ports Illustrating Synchronous Communication by Handshake Process.

The control lines used to communicate status information are sometimes called handshake lines. The process of controlling information transfer is called the handshake process.

Coordination is required between the microcomputer and the peripheral to prevent data loss. For example, if the peripheral sends new information before the microcomputer has read the temporary input register, the register can be altered to reflect the new information, erasing old data.

In the same manner, if the microcomputer has not received the go-ahead sign from the peripheral, and places new information in the temporary output register before the peripheral has read the previous information, the data is lost.

Thus, control lines protect the integrity of the data being transferred, and allow the transfer to occur at non-regular, or asynchronous, intervals.

In serial communication, start and stop bits before and after each byte, or word, provide special signals for the microcomputer that define data.

A device called a universal asynchronous receiver transmitter (UART) is basically a shift register that not only converts serial data into parallel data, but also provides bit-checking and parity-checking functions. The UART makes the rate at which data is fed into the microcomputer consistent and regular. It also transmits data from the microcomputer to an output peripheral at a consistent rate.

Consistent transmission rates are defined by using a baud rate. The term "baud" refers to the maximum number of signals that may be transmitted over a communication line and is a function of the duration of the shortest signal.

Baud is not identical with bits-per-second rate due to the fact that bits may be communicated by signals of varying lengths.

Synchronous Communication

In synchronous communication, data is transferred at a regular rate that is determined by a special control line, or external clock synchronization. Although the initial process of getting the transmitter and receiver in sync may require additional hardware and software, this method provides faster I/O communication – once synchronization has been established.

During synchronous communication, the microprocessor is performing an I/O operation at the same speed as the data transfer rate. For example, if data is being transferred at 1200 baud, the microprocessor is performing one I/O operation every one twelve-hundredths of a second.

As with UARTs, there are universal synchronous receiver transmitters (USRTs) available in integrated circuit form that perform approximately the same functions.

INTERRUPTS

Most microcomputers have the ability to temporarily halt, or interrupt, the normal execution of a program upon receipt of a signal from an external peripheral, and jump to a subroutine that deals with the request of the peripheral. Upon completion of the task requested by the peripheral (normally a data transfer), the original program resumes at the point it was originally interrupted.

Also called a program interrupt, this method of responding to I/O devices enables the microcomputer to perform operations until its attention is diverted to I/O. This differs from the regular sampling method where the computer is regularly looking for I/O data.

There are several types of interrupts. Polled interrupts utilize a single control channel from all I/O devices to the microprocessor. When an I/O device needs to transfer data, it places an interrupt request signal on the control channel. The computer temporarily halts execution of the existing program, storing its place so that it may return to the same point and resume operation. The computer then polls each I/O device to see which needs assistance. This is normally accomplished by having the peripheral place an interrupt request bit in the status register of its I/O port. The computer checks the various I/O status registers until it finds the device making the interrupt request.

There is usually an order of priority in which the microcomputer polls the I/O devices. Therefore, if two or more devices make interrupt requests at the same time, they are responded to in polling order. This order may be altered according to prearranged priority status codes.

In large computer systems connected to literally hundreds of I/O devices, this method for responding to interrupt requests can be time-consuming.

A faster method, called vectored interrupts, allows the computer to know immediately which device has made an interrupt request through certain signals. These signals are provided in several ways.

One method requires a separate line from each I/O device that causes a jump to a subroutine dealing with that specific device. However, this method is seldom used in microcomputers.

Most microcomputers make use of status bits, called flags, that signal when a program interrupt is requested. The address that corresponds to the specific I/O port, and thus to the device, is sent to the microprocessor, initiating the subroutine that corresponds to the device. In the case of multiple requests, devices called priority encoders send the addresses to the microprocessor in priority coding order.

Interrupts may be multiple level, meaning that an interrupt may be interrupted, and then that interrupt may be interrupted, and so forth. In each case, the microcomputer stores the address of the next instruction to be executed in a stack register in order to know where to resume a subroutine.

Interrupts may also be disabled in most microcomputers, meaning that under certain conditions critical to the software or hardware, the microcomputer will totally ignore interrupt requests.

I/O BUSING

Almost all microcomputers use the same data buses for I/O transfers as well as for memory operations, where data are being stored into, or read from, memory. As previously mentioned, I/O peripherals may even be treated as memory in such cases as disk and tape storage.

It is important that all data being routed back and forth over the same lines be carefully coordinated in order to be communicated properly. Several methods exist to provide this coordination. These methods follow:

1. Isolated I/O, where the address for memory and I/O data are handled separately in the decoding process.
2. Memory mapped I/O, where I/O ports are handled as locations in memory.
3. Attached, or incorporated I/O, where I/O ports are located in the microcomputer's CPU or memory sections within an integrated circuit.

Isolated I/O uses separate bus drivers, or devices that place the actual signals on the data bus. I/O signals often differ from memory signals, which seldom require handshaking or other control and coordinating signals. As a result, separating the I/O bus control from memory bus control has its advantages. Many times, programs are clearer. The physical design of circuitry that deals with I/O or memory may take advantage of this difference in the nature of the two forms of data. But these designs are usually less flexible, and require extra (or duplicate) instructions and decoding.

The biggest advantage of memory mapped I/O lies in the programming for microcomputers equipped with this feature. Program instructions can fetch data from a peripheral device as easily as from memory locations – without additional instructions required by other methods to physically get the information to the microcomputer from the peripheral. This makes programming easier, but can also make programming harder to understand if adequate documentation does not exist to distinguish where data originated.

Attached or incorporated I/O is utilized when the small physical size resulting from this method is a benefit. There are limitations in the number and complexity of I/O devices associated with this method. Integrated circuits containing the CPU, memory, and I/O control on one chip are more expensive.

I/O HARDWARE

The selection of I/O devices is widely-varied. A brief summary will be given for some more common devices.

Flip-flops have already been mentioned as one method for temporarily storing I/O data to help diminish the problems associated with wide variance in input/output rates.

A device known as a monostable multivibrator, or one-shot, transforms short pulse signals into longer, fixed-length signals. The one-shot multivibrator can smooth out peripheral signals to facilitate their interpretation by the microcomputer. These devices are variations of the bistable multivibrators discussed in Module MH-01, "Digital Components."

Shift registers have been mentioned as devices that can transform data from a serial to parallel form, and vice versa.

Counter circuits, also described in Module MH-01, can be used to drive I/O devices such as displays.

A selector chooses a single output from multiple inputs, while a decoder produces a single output based on multiple inputs. A third similar device (called a multiplexor) produces a single output (from multiple inputs) that contains all the information of each input.

All three devices have multiple inputs and a single output. However, there are similarities and distinctions between devices. A selector selects one of several inputs; a decoder produces an output based on the information provided by a coded input. A multiplexor integrates information from multiple channels onto a single channel.

SWITCHES

Just as the simple switch is the basis for complicated logic devices, it is also the most simple form of input and is the foundation for more complicated forms of input.

Push-button switches, such as those used in keyboards, are mechanical devices. A problem sometimes encountered with mechanical switches is called bounce. This phenomenon occurs when physical pressure exerted to close the switch produces Newton's famous equal and opposite reaction, causing the contact points to momentarily open again before stabilizing in the closed position.

This problem can be cured through software by sampling output (after a sufficient delay) to account for switch bounce. Hardware devices, such as the one-shot mentioned previously, help transform the two quick, successive ON/OFF states produced by switch bounce into a single, definite pulse.

Figure 3 illustrates a debouncing technique for a single pole, double throw (SPDT) switch, utilizing the same crossed-NAND configuration of Module MH-01.

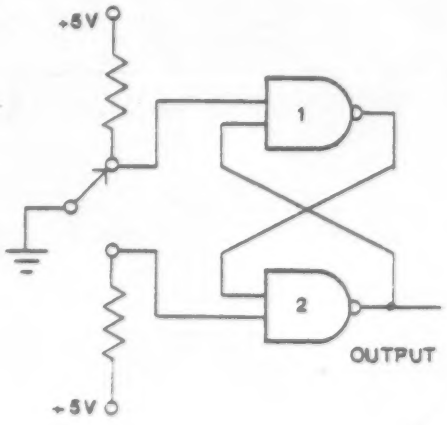


Figure 3. Using Crossed NAND Gates to Remove the Negative Effects of Switch Bounce.

A keyboard, like those used to input information into teletype machines or video display terminals (VDTs), is nothing more than a group of simple switches.

The word terminal, as used in video display terminal, implies an integrated I/O device that has the capability of sending information to the microcomputer and receiving information from the microcomputer.

Currently, VDTs use a cathode-ray tube (CRT) as the display device for information. Newer technologies involving flat-panel image devices may eventually replace CRTs.

CRT-based VDTs are currently the most popular and powerful form of I/O device because of their ability to display large quantities of information almost instantaneously. This information can be in color and in graph or table form. CRT-based VDTs can show movement or attract attention through blinking verbage or audio signaling. Interactive CRTs use light pens or even touch-sensitive screens to allow the user greater flexibility in information exchange.

The teletype is also a computer terminal that uses keyboard input, but display is normally accomplished through a typewriter-like impact printer.

DISPLAY

Printers are a basic form of computer output. They range from impact printers that strike a single character at a time to line printers that print an entire line of information simultaneously. A newer technology, called ink jet printing, manipulates a stream of ink droplets onto paper to form characters, thereby greatly decreasing printer

noise. Some high-speed printers can attain print rates of over 4,000 lines per minute.

Printers are useful storage mediums – generally of information for human consumption only. Computer data retrieval of printed material is a relatively slow process, accomplished through input devices called optical character readers (OCRs). More efficient forms of I/O storage/retrieval devices will be discussed shortly.

For output, the simple lamp functions similarly to the switch, in that it indicates information by means of a simple ON/OFF condition.

The analogy also holds that simple lamps are build up into more complex systems, just as switches formed the basis for keyboards. Light-emitting diodes or small incandescent bulbs can be used to form segmented displays and other video displays: information is still communicated through ON/OFF lighting.

In order to communicate with human beings, computers must transform our accepted symbolic communication systems, such as base 10 numerals and the 26-letter alphabet, into binary form, and vice versa. This is accomplished through various systems of coding.

I/O CODING

The binary coded decimal system, or BCD, is used extensively in microcomputers and especially in calculators.

Standard BCD uses the regular binary notation for the decimal digits 0-9, but uses a separate group of four binary digits to represent each digit in decimal notation.

For example, the decimal number 99 in standard binary form is 1100011, or seven binary digits. The representation for decimal 99 in BCD is two binary representations of decimal 9, or 10011001, requiring one more bit to communicate decimal 99 than ordinary binary coding.

The circuitry required to perform arithmetic operations in BCD is more complex than standard binary coding, but BCD codes can be interfaced to widely-available devices that convert BCD to numeric characters in a display, making the use of BCD coding or its variations in calculator circuitry very popular.

Also, since a 4-bit unit can represent any decimal digit, the code fits conveniently into the standard word lengths of most microcomputers (4-, 8-, or 16-bit words).

Alphanumeric codes must represent not only decimal numerals, but also the alphabet and any special characters required in communication.

A 5-bit code, called baudot, is used in some telegraph and teletype systems to represent alphanumeric characters in both upper and lower case.

EBCDIC, or extended binary coded decimal interchange code, is used in large computer systems, such as those made by IBM, to represent alphanumerics.

The most popular code for alphanumerics being used in microcomputers and their peripherals is called ASCII (pronounced AS-KEE) which stands for American Standard Code for Information Interchange. Table 1 contains a listing of ASCII coding.

ASCII uses seven bits to represent 128 characters. The eighth bit is used for parity checking.

TABLE 1. AMERICAN STANDARD CODE FOR INFORMATION INTERCHANGE (ASCII).

ASCII Code	Character	ASCII Code	Character	ASCII Code	Character
00	NUL	2B	+	56	V
01	SOH	2C	,	57	W
02	STX	2D	-	58	X
03	ETX	2E	.	59	Y
04	EOT	2F	/	5A	Z
05	ENQ	30	0	5B	[
06	ACK	31	1	5C	\
07	BEL	32	2	5D	]
08	BS	33	3	5E	^(↑)
09	HT	34	4	5F	_(←)
0A	LF	35	5	60	`
0B	VT	36	6	61	a
0C	FF	37	7	62	b
0D	CR	38	8	63	c
0E	SO	39	9	64	d
0F	SI	3A	:	65	e
10	DLE	3B	;	66	f
11	DC1(X-ON)	3C	<	67	g
12	DC2(TAPE)	3D	=	68	h
13	DC3(X-OFF)	3E	>	69	i
14	DC4	3F	?	6A	j
15	NAK	40	@	6B	k
16	SYN	41	A	6C	l
17	ETB	42	B	6D	m
18	CAN	43	C	6E	n
19	EM	44	D	6F	o
1A	SUB	45	E	70	p
1B	ESC	46	F	71	q
1C	FS	47	G	72	r
1D	GS	48	H	73	s
1E	RS	49	I	74	t
1F	US	4A	J	75	u
20	SP	4B	K	76	v
21	!	4C	L	77	w
22	"	4D	M	78	x
23	#	4E	N	79	y
24	$	4F	O	7A	z
25	%	50	P	7B	{
26	&	51	Q	7C	|
27	'	52	R	7D	}(ALT MODE)
28	(	53	S	7E	~
29	)	54	T	7F	DEL(RUB OUT)
2A	*	55	U		

The most widely-used numeric display consists of LEDs, or light-emitting diodes. There are normally seven distinct segments that form numerals and some letters, when lit in various combinations. A special code, called seven-segment code, represents these various lighting combinations used to represent many alphanumeric characters. A portion of this code is listed in Table 2.

TABLE 2. SEVEN-SEGMENT CODE.

Decimal Digit	Seven-Segment Code (hexadecimal)
0	3F
1	06
2	5B
3	4F
4	66
5	6D
6	7C or 7D
7	07
8	7F
9	6F or 67
Uppercase Letter	**Seven-Segment Code (hexadecimal)**
A	77
C	39
E	79
F	71
H	76
J	1E
L	38
O	3F
P	73
U	3E
Y	6E

I/O STORAGE DEVICES

One of the earliest I/O devices constructed for the primary purpose of storing digital information, punched paper tape is still being used in microcomputer applications (from teletypes to typesetters) to read and store programs and data.

Punched paper tape is limited in speed and slightly bothersome in its physical storage requirements. Furthermore, paper tape is sequential in nature, meaning that a spool of tapes must be physically wound to the end of the spool to retrieve data stored in that location.

The actual data are stored by using various coding systems to punch combinations of holes into sequential positions on the paper tape. These holes are read by passing the tape through a reader that uses light or mechanical contact to sense the presence or absence of holes.

A similar system, using punched holes in cards rather than tape, allows information to be physically sorted, or rearranged to a particular order, to facilitate information retrieval.

Magnetic tape is an improvement over punched tape and cards, both in speed and storage characteristics.

However, magnetic tape is still sequential in nature, whether it is a simple modified audio cassette player, such as the one used by the KIM-1 microcomputer, or large, high-speed tape drives used in the largest computer systems.

Magnetic disk systems maintain the advantages of magnetic storage but have added a random-access capability that greatly increases the speed of information storage and retrieval.

These devices use disks with small units of magnetically alterable material that spin at high speeds, similar to a high-speed phonograph record composed of audio tape material. A read/write head, or encoding/decoding device, can move freely along a radius of the disk, enabling the device to quickly jump from the innermost track to the outermost track, or any location in between.

Large systems use multiple layers of hard disks to increase storage capacity. A simplified diagram of multiple-layer disk is shown in Figure 4.

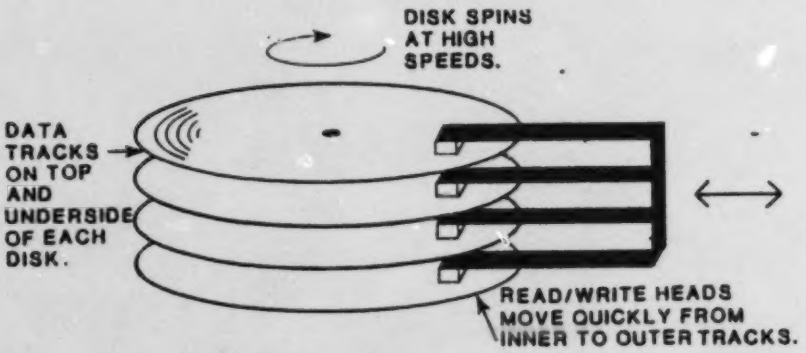

DISK SPINS AT HIGH SPEEDS.
DATA TRACKS ON TOP AND UNDERSIDE OF EACH DISK.
READ/WRITE HEADS MOVE QUICKLY FROM INNER TO OUTER TRACKS.

Figure 4. Multi-Layer Magnetic Disk Storage/Retrieval System.

Floppy disks, so-named because the material comprising the magnetic disk is flexible, are widely used in microcomputer systems. Lightweight, durable, and relatively inex-

pensive, this form of I/O storage can access information from any location on the spinning disk in a few thousandths of a second. Some floppy disks can store over 8 million bits of information.

The same principles of random-access disk storage are utilized in sealed configurations of hard disks, called fixed disks, which are more reliable, much faster, and have a greater storage capacity than floppy disks.

FUTURE DEVELOPMENTS

Rapid strides are being made in speech recognition and synthesis devices as a result of their growing popularity in consumer electronic devices. It is not known to what extent the ability to talk to and listen to a microcomputer will alter the nature of current I/O communication.

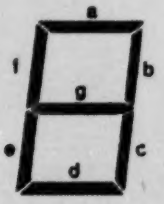
As solid-state integration processes become more advanced, more and more I/O control functions are being included physically on the microprocessor chips.

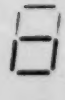
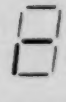
As the price of these chips drops, more and more smart peripherals appear. Smart peripherals use a dedicated, rather than general purpose, microprocessor. The microprocessor will have only one purpose, such as controlling a VDT. This form of dedicated use is also called distributed processing.

The trend toward distributed processing and smart peripherals extends not only to the devices mentioned in this module, but also to instruments, sensors, and special devices that may be encountered in the field by an energy technician.

EXERCISES

1. Identify the following abbreviations, and define the terms for which they stand:
   a. UART.
   b. USRT.
   c. CRT.
   d. VDT.
   e. OCR.
   f. BCD.
   g. EBCDIC.
   h. ASCII.
2. Table 2 presented a portion of the standard seven-segment code. In the figure below, the corresponding positions of the display have been marked to show the binary number that controls each segment. In the spaces provided, fill in the binary code that would be required for the letters that can be made using seven-segment code. Then convert the binary code to hex. (The first two codes are from Table 2. Use it to check your answers for the first two.)



	Readout	Binary Code (MSB is 0)	g	f	e	d	c	b	a	Hex Code	
1.		0									
2.		0									
3.		0									
4.		0									
5.		0									
6.		0									
7.		0									
8.		0									
9.		0									
10.		0									

This laboratory takes a closer look at the KIM-1 microcomputer in terms of the way it inputs and outputs information.

The figure below is a rough diagram of the basic layout of the KIM-1. Notice that many of the devices that make up the computer are related to I/O functions.

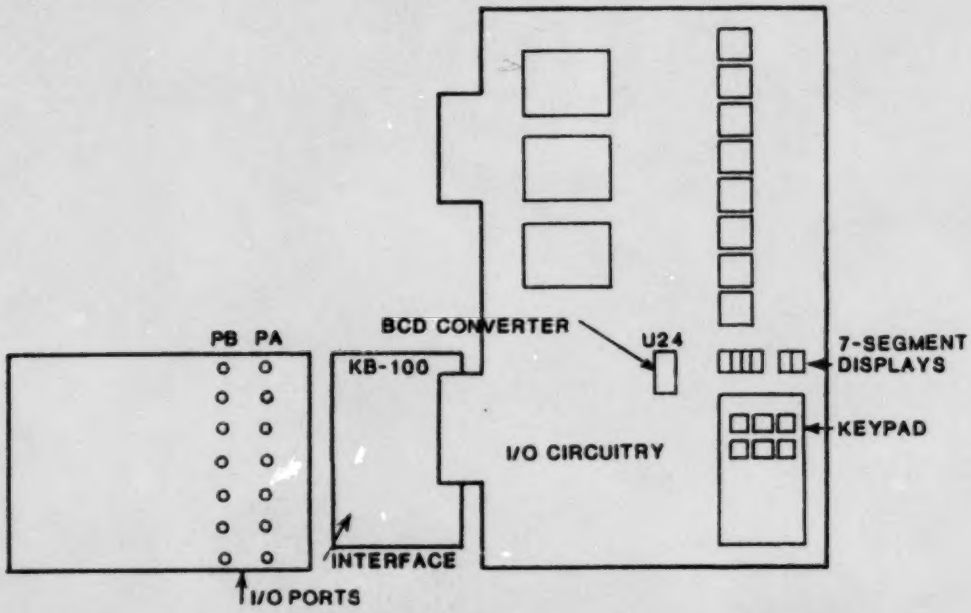


Figure 5. KIM-1 Microcomputer Layout.

Notice that the I/C chip labeled U24 is a BCD decoder/driver, and that its output is connected to the resistors and transistors located above the seven-segment displays, which serve as the driving circuitry for the LED display.

Notice that the LEDs do, in fact, have seven segments. Using the keypad as an input keyboard, the student may view

each of the 16 characters that are possible on these seven-segment displays.

The primary input/output ports of the KIM-1 are two bytes located at addresses 1700 (hex) and 1702 (hex). When defined as input, the bits at these locations result from logic values being applied on any of the 16 lines by means of input voltages. When defined as output, these lines will contain the logic voltages reflected at these locations in the data fields.

When in the input mode, the peripheral output buffers are in the 1 state. A pull-up resistor acts as less than 1 TTL load to the peripheral data lines. Therefore, all lines reflect logic 1 unless grounded.

Turn the power to the KIM-1 on. Press AD and 1700 to activate pins PA-0 through PA-7 as input ports. The data display should now read FF. Why?

Ground the switches according to Data Table 1 and record the result of the data display.

When ports are set at output, voltages on the ports reflect data that has been stored at the dedicated address.

IMPORTANT: REMOVE ALL GROUNDS FROM PORTS BEFORE CHANGING THEM TO OUTPUT, SO THAT AN OUTPUT VOLTAGE WILL NOT BE CONNECTED DIRECTLY TO GROUND.

Store FF at address 1701. This will convert port A (PA) to an output port. Enter the data values into address 1700 and record the corresponding voltage levels in Data Table 2.

DATA TABLE 1: RESULT OF THE DATA DISPLAY.

Ground Pins								Data Display	
PA–0	1	2	3	4	5	6	7		
1. G	G	G	G	G	G	G	G	____	____
2. G		G		G		G	G	____	____
3.	G		G		G	G		____	____
4.			G				G	____	____
5. G	G	G	G					____	____
6.				G	G	G	G	____	____
7.			G	G				____	____
8. G	G					G	G	____	____

DATA TABLE 2: CORRESPONDING VOLTAGE LEVELS.

Data Input	Voltage Level						
	PA–0	2	3	4	5	6	7
1. AA	—	—	—	—	—	—	—
2. F1	—	—	—	—	—	—	—
3. 88	—	—	—	—	—	—	—
4. 7C	—	—	—	—	—	—	—
5. E3	—	—	—	—	—	—	—
6. 5B	—	—	—	—	—	—	—
7. 62	—	—	—	—	—	—	—
8. 9D	—	—	—	—	—	—	—

REFERENCES

Bartee, Thomas C. Digital Computer Fundamentals. New York: McGraw-Hill, 1977.

Gilmore, Charles M. Beginner's Guide to Microprocessors. Blue Ridge Summit, PA: Tab Books, 1977.

Kershaw, John D. Digital Electronics: Logic and Systems. North Scituate, MA: Duxbury Press, 1976.

KIM-1 User Manual. Norristown, PA: MOS Technology, Inc., 1976.

Leventhal, Lance A. Introduction to Microprocessors: Software, Hardware, Programming. Englewood Cliffs, NJ: Prentice-Hall, 1978.

Soucek, Branko. Microprocessors and Microcomputers. New York: John Wiley & Sons, 1976.

Texas Instruments Learning Center. Understanding Solid-State Electronics. Dallas, TX: Texas Instruments, Inc., 1978.

Tinker, Robert F. Microcomputers. Cambridge, MA: Technical Education Research Center, 1978.

Wise, Kensall D. et al. Microcomputers - A Technology Forecast and Assessment to the Year 2000. New York: John Wiley and Sons, 1980.

TEST

1. Would a temperature sensor connected as an input device to a microcomputer use a serial or parallel input? Why?

2. Explain briefly the handshake operation. Discuss the function of each control link.

3. Why is the concept of interrupts useful to computer I/O operations?

4. Compare the advantages and disadvantages of memory mapped and isolated I/O.

5. Define the following terms:
   a. Selector.

   b. Decoder.

   c. Multiplexor.

6. What distinguishes a terminal from other I/O devices?

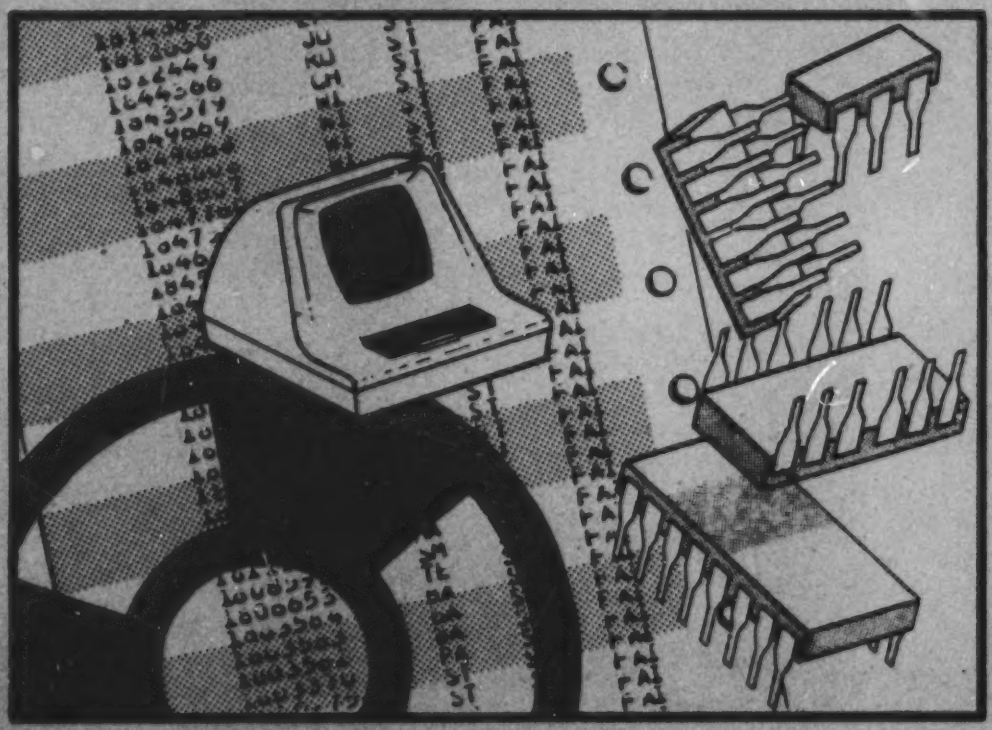
7. Briefly explain the concept of binary coded decimal.

8. What is the advantage of a floppy disk over punched paper tape?

ENERGY TECHNOLOGY

CONSERVATION AND USE

MICROCOMPUTER HARDWARE



MODULE MH-04

ANALOG/DIGITAL CONVERSION

TECHNICAL EDUCATION RESEARCH CENTER - SOUTHWEST
4800 LAKEWOOD DRIVE, SUITE 5
WACO, TEXAS 76710

INTRODUCTION

Many microcomputer applications involve analog input and output signals. Analog signals cannot be processed directly by digital computers, but must first be converted to a digital form (by using an analog-to-digital converter) before they can be fed into a computer. Similarly, a computer cannot generate analog outputs directly, but must use digital output signals to control digital-to-analog converters. This module introduces some of the more common analog/digital converters and conversion techniques.

PREREQUISITES

The student should have completed MH-01, "Digital Compoments"; MH-02, "Semi-Conductor Logic Families"; and MH-03, "Input/Output Devices and Techniques."

OBJECTIVES

Upon completion of this module, the student should be able to:

1. Construct a resistor-ladder digital-to-analog converter.
2. Describe the operating principles of three types of analog-to-digital converters.
3. Define the following terms:
   a. Linear.
   b. Linearity.
   c. Bandwidth.
   d. Resolution.
   e. Precision.

f. Accuracy.
g. Shaft encoder.
h. Transducer.
i. Resistor ladder.
j. Weighted summation.
k. Multiplying DAC.
l. Op-amp.
m. Analog comparator.
n. Quantizing error.
o. Successive approximation ADC.
p. Counter-comparator ADC.
q. Simultaneous ADC.
r. Reference voltage.

ANALOG/DIGITAL CONVERSION

As stated in previous modules, digital signals communicate information by means of separate, distinct states, while analog signals vary in current, voltage, or resistance in order to communicate information.

Theoretically, analog signals may represent an infinite number of values. In the first place, there are as yet no known limits, either high or low, for such electrical properties as voltage and current. Secondly, just as there are an infinite number of points which make up a line in geometry, there are an infinite number of single values that compose a gradually increasing or decreasing signal.

Figure 1 represents an analog signal that varies from 0 volts to +5 volts and back again to 0 volts. This variance is drawn over a grid that divides the signal into 1/2-volt

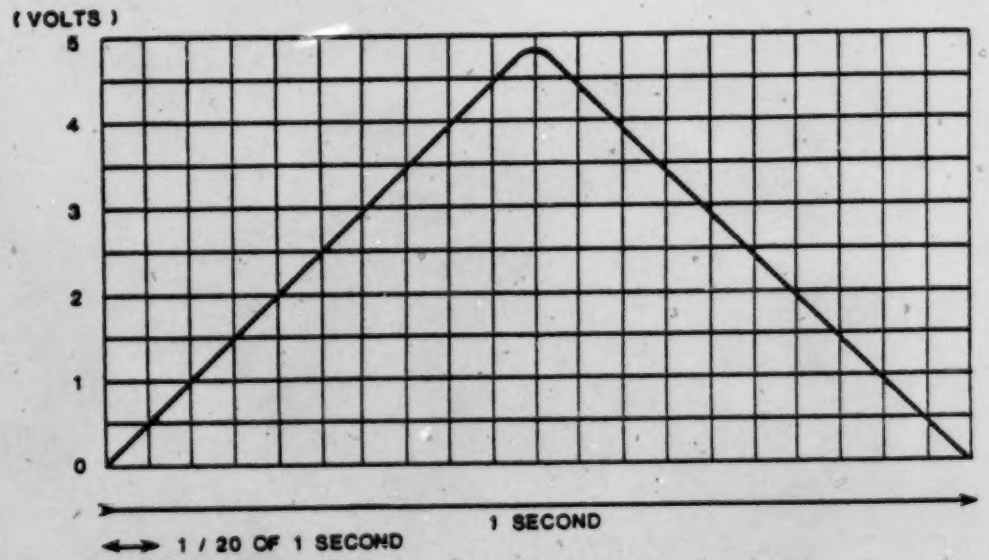


Figure 1. Graph of a Linear (Analog) Signal.

increments vertically, and 1/20th of a second increments horizontally. Notice the smooth variance in the analog signal. This is the reason analog signals are also called linear signals.

This waveform could be a close representation (analogy) of a real event in the physical world, in the sense that most real world events vary gradually between parameters rather than step up or down in precisely-defined increments. For example, the waveform might represent outside temperature from the time the sun appeared over the horizon until after sunset.

Suppose that each 1/2-volt increment symbolized a single unit of measurement. Then, in the space of one second, this signal has, in effect, counted from 0 to 10 and back to 0.

But what if each 1/50th of a volt represents one unit? As mentioned before, in theory the exact number of possible divisions is infinite. Dividing each volt into 50 units will allow this same signal to, in effect, count from 0 to 250 and back to 0.

In digital form, the decimal digits 0 through 10 are limited to binary representations; it takes a minimum of four binary digits to represent each decimal digit. Therefore, in order to count from 0 to 10 and back to 0 in digital form, as was done in the first analog example, the second would need to be divided into a minimum of 80 parts rather than 20. This is illustrated in Figure 2.

In order to count from 0 to 250 and back to 0, a digital signal requires a second to be divided into a minimum of 2,000 parts, as opposed to the 500 needed for the analog signal.

The number of divisions required to send information by either digital or analog means is called frequency bandwidth, or simply bandwidth.

Figure 2. Binary (Digital) Count From 1 to 10 and Back to 0.

Although in the previous figures the bandwidth required for an analog signal was only 25% of the bandwidth required to communicate the same information by digital means, the ratio can be even more heavily weighted in favor of analog efficiency when comparing signals required for television (video) transmission.

And yet, each method of communication has its own advantages and disadvantages. Therefore, conversion from digital to analog or analog to digital is necessary whenever circumstances favor the advantages of one over the other.

Most of the real world is composed of analog information. Information important to energy consumption and conservation, such as temperature, power consumption, or air and liquid flow measurements, are analog. In order to take advantage of the ever-increasing cost effectiveness of digital microcomputer information processing and control, the student must necessarily have an understanding of conversion methods between the two communications methods.

The first figures also illustrate two other important characteristics used in analog and digital conversions. It was shown that by varying the rate at which an analog signal is sampled (in other words, the number of time divisions, or counts), the same analog signal can yield 10 units or 250 units.

Obviously, a signal described by 250 digital equivalents will be more precise than the same signal described by 10 digital equivalents.

The greater the number of binary representations of an analog signal, the greater its resolution and precision. The actual difference in the measured value of an analog signal at a given time and its digital equivalent yields the accuracy of the conversion.

In this module, the actual devices that perform the conversions will be abbreviated DAC, for digital-to-analog coverter, and ADC, for analog-to-digital convertor. Whenever the technique of conversion is being referred to, it will be abbreviated a-d conversion or d-a conversion.

Both a-d and d-a conversions can be made by mechanical or electrical means. After a brief discussion of mechanical conversion devices, the majority of this module will deal with electrical conversion techniques.

MECHANICAL CONVERSION DEVICES

In many cases, a digital representation might be required to indicate the position of a potentiometer, or rotary switch. Such a switch might be found on the front panel of a washing machine to indicate cycles, for example. A device called a shaft encoder (by means of physical contact to conduct electricity or photoelectric sensing) can perform an a-d conversion.

Figure 3 illustrates the basic concept. The darker areas represent conductive material (in the case of contact devices) or transparent material (in the case of photoelectric devices), while the light areas represent nonconducting or opaque areas.

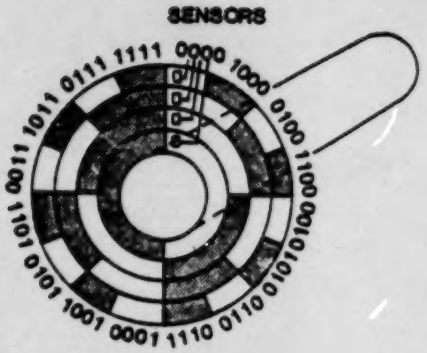


Figure 3. Shaft Encoder Used
For Mechanical a-d Conversion.

As the shaft rotates, the disk also rotates, and the binary code indicated on the sensors represents the position of the switch.

Mechanical converters such as these have their difficulties, such as when the position of the disk is such that the sensors are on the line, resulting in an ambiguous digital signal. Various codes, software, or multiple position sensor techniques have been developed to overcome such problems. Other problems include the wear associated with friction in contact devices, and dust or other contaminants in photoelectric devices.

ELECTRICAL CONVERSION TECHNIQUES

The trend today is toward transducers, which sense analog conditions in the real world, convert them to electrical signals, and can then be converted to digital form by electrical, rather than mechanical means.

Solid-state electronic devices such as phototransistors actually produce analog voltages in response to light, which can be measured by means of an ADC.

There are even solid-state devices (such as the National AN-132 Temperature Transducer) that directly sense temperature and convert the information to electrical signals. This same device can also be used to detect movement in air or liquid, to measure wind velocity, or to sense position.

DIGITAL-TO-ANALOG COMPUTER

A very simple DAC can be formed by a network of resistors that are interconnected so that opening and closing binary switches produces a varying (analog) voltage output.

Figure 4 represents a 4-bit resistor network DAC.

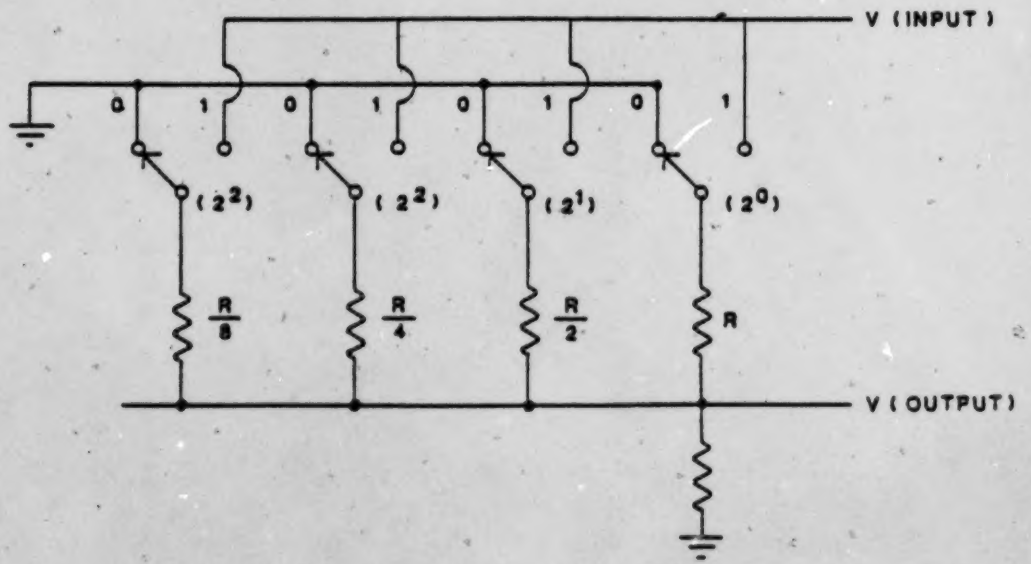


Figure 4. DAC Formed by a Network of Resistors.

The resistance value selected for R determines values for the remaining resistors. In a circuit, the switches are replaced by transistors controlled by the output of a storage register or counter circuit.

A slightly different circuit design, known as a resistor-ladder network, allows the same d-a conversion characteristics, but may be formed with only two values of resistor (or one value, if additional series circuitry is employed).

Figure 5 illustrates a resistor-ladder network.

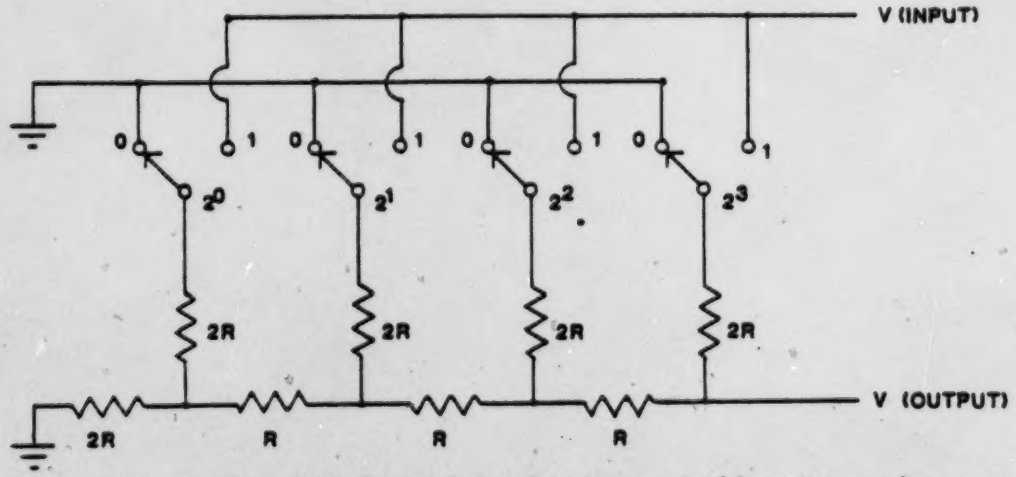


Figure 5. DAC Formed by Resistor-Ladder Network.

This resistor ladder could be expanded easily by adding additional switches and resistors in a similar fashion.

Switches representing increasing binary values access a voltage path of increasing resistance. Combinations of closed switches set up series and parallel circuits that produce an analog representation of the switches at the voltage output.

This method of analog generation is sometimes called weighted summation, because it progressively adds the resistive values in order to vary the analog value produced at the output of the DAC.

A detailed investigation into the characteristics of a resistor-ladder network make up the laboratory for this module.

If the switches of the DAC in Figure 5 were controlled by the output of a binary counter that generated the binary equivalents of decimal 0 through 9 and back to 0, a waveform similar to the one in Figure 6 would result.

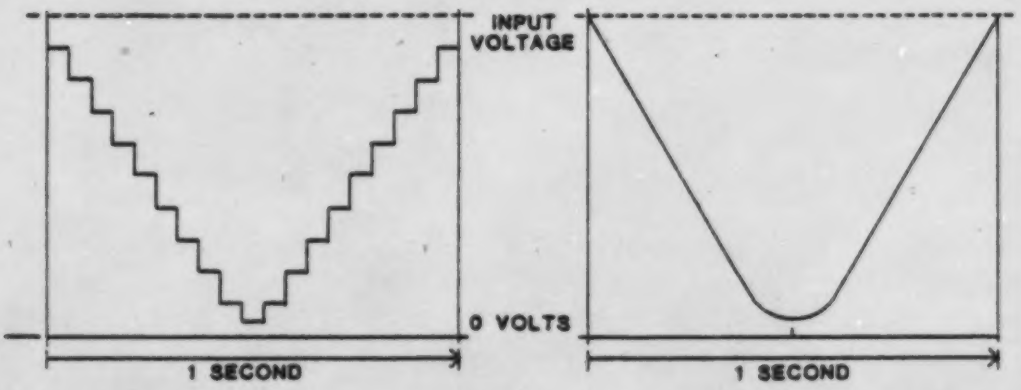


Figure 6. Comparison of DAC and Pure Analog Waveforms.

Figure 6 also compares this DAC waveform with an equivalent pure analog waveform. Notice the similarities.

In this particular case, setting the binary switches to the lowest equivalent decimal value produced the highest voltage, while each increment upward in decimal value resulted in a corresponding decrease in voltage. The output of this

circuit could be inverted so that the most positive voltages represented the most positive decimal equivalents, as in Figure 1.

Although seemingly simple in design, resistor network DACs require precision power supplies, transistors, and other circuit components to insure adequate accuracy for almost any conversion requirement.

Voltage generators that produce voltage inputs to the resistor-ladder network must be very precise due to the loading of the resistor network, and these generators must respond to logic level control (digital inputs).

Voltage generators sometimes use an analog reference voltage for comparison to insure precision. Some DACs, called multiplying DACs, allow this analog reference to be varied, altering the output of the device to the product of the analog voltage times the digital input.

This gives multiplying DACs a variable range. Range is specified for all DACs and pinpoints the minimum and maximum analog voltage values capable of being produced by the device.

Resistor networks that divide voltages are one type of DAC technique. Current may also be used to communicate analog information, using a transistor network.

Although higher conversion speeds may be attained, the voltage range of these devices is much lower than the voltage level being applied to the DAC.

The analog inputs to an ADC must also be specified. For both types of convertors, input and output ranges – as well as operating frequencies – are important characteristics in choosing the proper device for the application.

Today, large-scale integration techniques are reducing the physical size of both DACs and ADCs, with some conversion circuits being incorporated into the microprocessors themselves.

This integration is made possible by utilizing special types of transistor circuits, called operational amplifiers, that are well-suited for integrated circuit applications. Operational amplifiers may be substituted for the resistor networks that perform d-a conversions.

Operational amplifiers, or op-amps, are made from linear transistors, as opposed to the switching type transistors described in module MH-02, "Semi-Conductor Logic Families."

Two qualities of linear transistors make them well-suited for use in place of resistor networks. First, they have the ability to generate a varying (analog) output. Secondly, this can be done in response to multiple inputs which interrelate in a summing manner, similar to the weighted summation characteristic of a resistor network DAC.

Since the output characteristics of an op-amp can be expressed in mathmatical terms that coincide with the polynomial expansion of a binary number, they can be used very effectively to perform d-a conversions.

Capable of addition, the high gain characteristics of op-amps allow them to perform more elaborate functions, such as a convertor from BCD code in digital form to analog decimal form, by varying the gain characteristics of the devices used to compose the circuit.

ANALOG-TO-DIGITAL CONVERSION

All of the techniques for a-d conversion are basically trial and error approaches, where the unknown analog quantity is compared to a trial (or reference) quantity – which is then altered based on whether the unknown value was judged to be higher or lower than the reference.

Therefore, the key component in ADC circuitry is a device called an analog comparator. The logic symbol for this device is pictured in Figure 7 below.

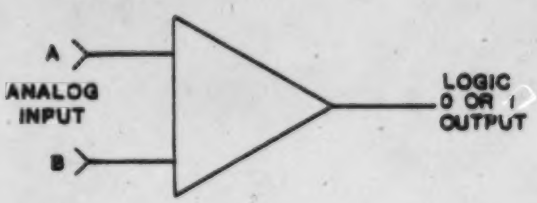


Figure 7. Logic Symbol for an Analog Comparator.

The function of this device is to compare two input voltages, here represented by A and B. If A > B, then the output of the device is logic 1. If B > A, then the output of the device is logic 0. If circuit design calls for reversed logic values to indicate a desired condition, the leads may be reversed. Thus, output is also reversed.

The actual circuitry that comprises an analog comparator is composed of resistors and bipolar transistors. Figure 8 is a schematic for a typical bipolar analog comparator.

When voltage A is greater than voltage B, the current through R_1 is greater than that through R_2. In this case, the emitter of the trigger transistor T_3 will be more positive than its base, and no current will flow through T_3. As a result, current will not flow through T_4 as well. Therefore, the output from the circuit will be logic 0.

When voltage B is greater than voltage A, the current through R_2 is greater than the current through R_1, turning on T_3, which in turn allows current to pass through R_4. The output of the circuit is logic 1.

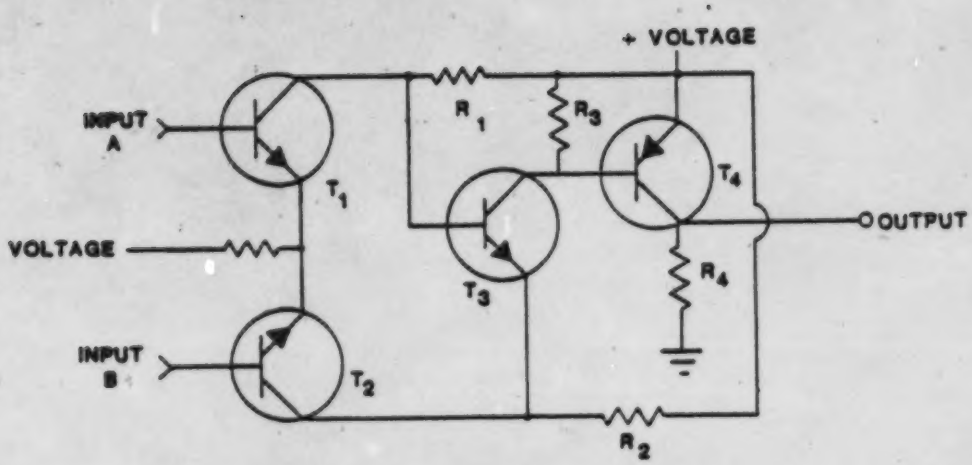


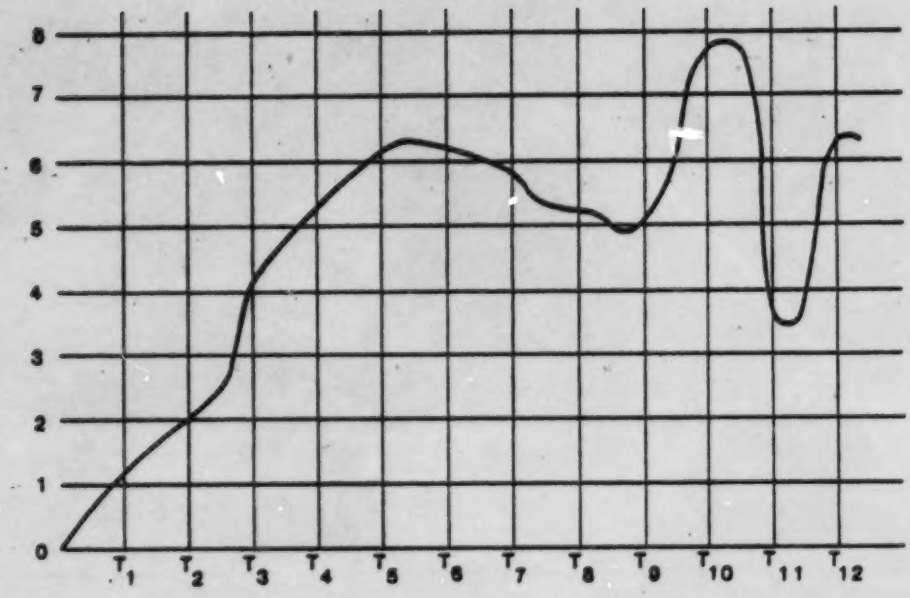
Figure 8. Analog Comparator Comprised of Bipolar Transistors.

RESOLUTION

Figure 9 graphs an analog waveform through 12 time divisions as the signal varies from 0 to 8 volts. This represents a resolution of one volt per count. If each volt represents a binary number, an increase of 1 volt increases the binary count by one. A decrease of one volt decreases the binary count by one.

The values listed for each count (T) could be the result of an ADC. Notice that even though the analog voltage varied from 6.25 volts at T_6 down to 4.9 volts between T_8 and T_9, the readings for T_7 through T_9 did not change.

This illustrates the concept of resolution as it relates to accuracy. Since at time T_6, the binary result could have been either 0110, or 0111, the accuracy of this ADC is limited



T_1	0001
T_2	0010
T_3	0100
T_4	0101
T_5	0110
T_6	0110
T_7	0101
T_8	0101
T_9	0101
T_{10}	0111
T_{11}	0011
T_{12}	0110

Figure 9.
ADC Output Illustrating
the Concept of
Resolution.

to +/- 1/2 the least significant bit. This accuracy rating is known as quantizing error, and is one element affecting the precision of an ADC.

If a straight line were drawn between the minimum and maximum input levels of a converter and then the maximum possible deviation from that straight line were plotted, the resulting graph would represent the linearity of the converter.

Linearity may be expressed as a fraction of the voltage equivalent to the least significant bit. In the case of Figure 9, the linearity of the converter would be +/- 1/2-volt.

The greater the speed of the converter, the greater the quantity of samples that may be taken of an unknown voltage in a given time period. Therefore, a greater resolution can be acheived.

There are three basic types of a-d conversion techniques: successive approximation, simultaneous conversion, and counter-comparator. Each of these techniques varies in the rate at which data can be converted.

COUNTER-COMPARATOR

The simplest method of a-d conversion uses a DAC (such as a resistor-ladder network), an analog comparator, and a binary counter.

Figure 10 diagrams a 4-bit counter-comparator ADC in its simplest form.

Initially, the counter is reset to 0 and begins counting UP. As each new count is input from the counter to the DAC, an increasingly higher voltage is compared with the unknown analog value.

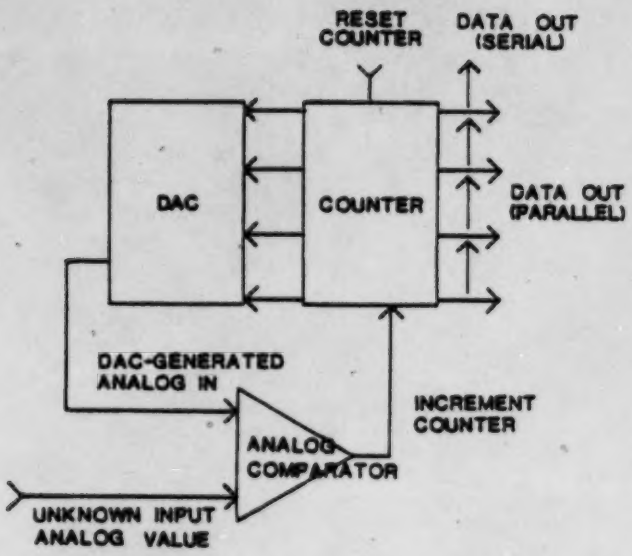


Figure 10. Block Diagram of Counter-Comparator ADC.

When the output of the analog comparator signals that the output voltage of the DAC has exceeded that of the unknown analog value, the value in the counter can be read out as the approximate equivalent to the unknown analog voltage.

This data may be read out in a parallel or serial fashion.

The counter-comparator method is slow because it begins at the minimum possible value for any unknown and increments in single units until an equivalent value is found.

An improvement on the simple counter-comparator takes advantage of a three-way comparator and an UP-DOWN counter. This system, diagramed in Figure 11, can start at an estimated middle ground for the unknown value and be guided toward the correct conversion by continually responding to a HIGH/LOW condition. If the estimated middle-value starting point is close to or above the actual middle value of the unknown quantity, the speed of the conversion can be doubled or more.

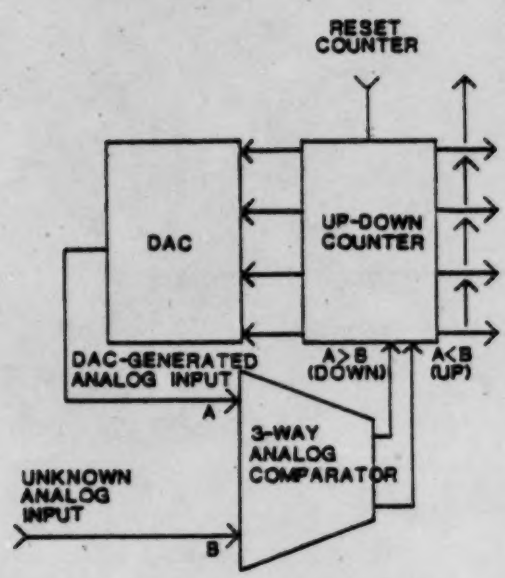


Figure 11. Block Diagram of an Improved Counter-Comparator ADC Using a 3-Way Analog Comparator and an UP/DOWN Counter.

SUCCESSIVE APPROXIMATION

The successive approximation method of a-d conversion takes advantage of the fact that each bit, from the most significant to the least significant, successively halves the voltage output of a DAC.

For example, in a 4-bit DAC, the maximum output analog voltage value would be 15 volts if each count represents 1 volt. If the MSB is 1, the output of the DAC is 8 volts, or approximately half its total possible output. Table 1 shows approximate values and their relationships as each successive bit represents a logic one.

In this method, the ADC circuit halves the total potential variance left in the unknown analog quantity with each

TABLE 1. APPROXIMATE VALUE OF BITS.

Bit Progression	Voltage Output	Approximate Fraction of Total Output
1000	8 volts	1/2
0100	4 volts	1/4
0010	2 volts	1/8
0001	1 volt	1/16

count. This means that each count produces a progressively closer approximation to the analog value being converted.

Once again, a trial and error conversion method has been used against a controlled reference value. However, with this method, speed is attained by moving UP or DOWN by half the remaining difference – as opposed to moving UP or DOWN incrementally toward the unknown value as in the improved counter-comparator system.

SIMULTANEOUS CONVERSION

Simultaneous conversion requires a separate analog comparator for each count in the digital output. In other words, a 4-bit ADC can represent 16 different counts (0-15) and would require 15 separate analog comparators. (A comparator is not required for the 0 count.)

When an unknown analog voltage is applied, every comparator with a reference voltage input BELOW the unknown value

will produce a binary 1 at its output, while those with reference voltage inputs HIGHER than the unknown will produce a binary 0.

The reference input to each comparator flows through a resistor network similar to those used in DACs in order to weight the reference voltage to properly correspond with the count position of each comparator.

In this manner, an a-d conversion can take place during a single clock period, but the large number of circuit components required for this method of conversion make it too expensive for most applications.

Although the simultaneous ADC is the fastest of the three methods, the successive approximation method is the more popular because it is less expensive, but maintains speed and accuracy.

MICROPROCESSOR-CONTROLLED CONVERSIONS

By taking advantage of software techniques and using the microprocessor itself for control of the conversion process, the amount of external circuitry and connections can be reduced.

This is especially the case when multiple analog inputs to the microcomputer require conversion at the same time.

The simplified block diagram in Figure 12 shows how the computer can control inputs (and thus the output) of a DAC in order to perform a parallel comparison of four different unknown analog inputs.

The microcomputer can also use various software schemes to generate trial voltages, such as the counting techniques or successive approximation techniques discussed.

However, the use of one DAC to provide trial voltages for multiple channels – which are then compared by the microcomputer – saves time and reduces the number of circuits required.

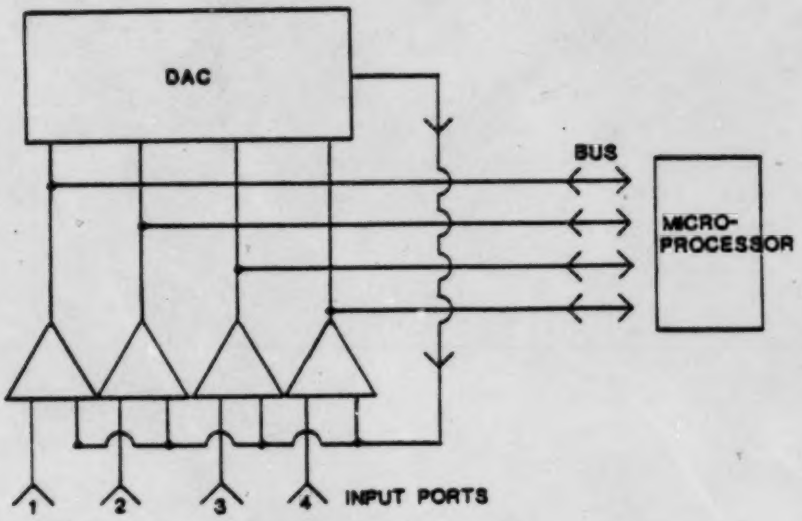


Figure 12. Simplified Block Diagram for Computer-Assisted Conversion.

EXERCISES

1. What digital bandwidth would be required to count from 1 to 3 and back to 1 a total of 5 times in one second?
2. The output of the analog comparator below is 1 if A > B, 0 if B > A. What would the output be under the following input conditions?

A	B	Output
+5	-3	_____
-3	0	_____
0	-1	_____
+6	+5	_____
-2	-1	_____

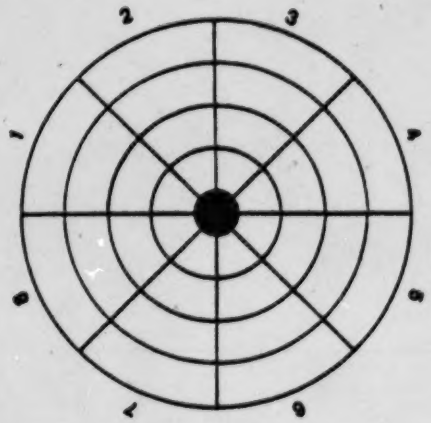
3. If the output range of a 3-bit multiplying DAC equaled one volt per count, what would the output range be if the reference voltage were doubled?
4. If the analog comparator used in Exercise 2 were driving an UP/DOWN counter in a counter-comparator ADC, what would the signal to the counter be, count UP or count DOWN?

A	B	Count UP or DOWN
+12	+11	_____
-2	-1	_____
-1	0	_____
0	-2	_____
-2	+2	_____

5. If one count equals one volt and there are 8 inputs to the DAC (a one-byte input), what would the range of the DAC need to be to take advantage of all possible input combinations?

6. If the circle below represents the disk of a shaft encoder, and dark areas are conductors, how would the disk look in order to represent binary equivalents of the following decimal numbers in the positions indicated?

1. 8
2. 2
3. 5
4. 7
5. 3
6. 1
7. 4
8. 6



LABORATORY MATERIALS

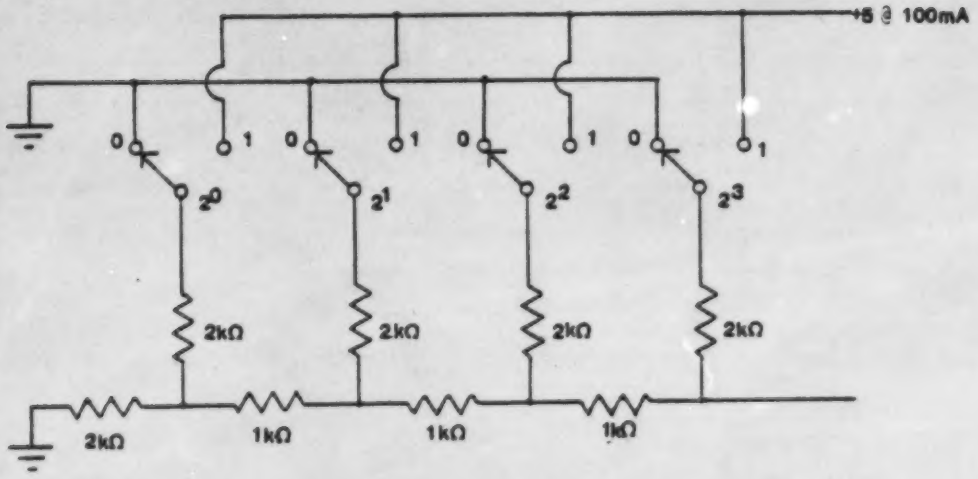
10 1-k Ω resistors.
3 2-k Ω resistors.
VOM.
Power supply generating +5 volts @ 100 mA.

LABORATORY PROCEDURES

The resistor-ladder network of Figure 5 is shown (on next page) with 1-k and 2-k Ω resistors used as values.

Construct this circuit and apply power at +5 volts, 100 mA. Use points A and B as test points for a VOM, and record the readings for current, voltage, and resistance as each

switch is closed. (Represent digital input.) Use the 25 mA scale for current readings, and the 10 DCV scale for voltage.



DATA TABLES

DATA TABLE 1. RECORD LAB RESULTS HERE.

Switch Positions	Voltage	Current	Resistance
1. 0000			
2. 0001			
3. 0010			
4. 0011			
5. 0100			
6. 0101			
7. 0110			
8. 0111			
9. 1000			
10. 1001			
11. 1010			
12. 1011			
13. 1100			
14. 1101			
15. 1111			

REFERENCES

Bartee, Thomas C. Digital Computer Fundamentals. NY: McGraw-Hill, 1977.

Givone, Donald D. and Roesser, Robert P. Microprocessors/ Microcomputers: An Introduction. NY: McGraw-Hill, 1980.

Kershaw, John D. Digital Electronics: Logic and Systems. North Scituate, MA: Duxbury Press, 1976.

Lenk, John D. Handbook of Microprocessors, Microcomputers, and Minicomputers. Englewood Cliffs, NJ: Prentice-Hall, Inc., 1979.

Leventhal, Lance A. Introduction to Microprocessors: Software, Hardware, Programming. Englewood Cliffs, NJ: Prentice-Hall, Inc., 1978.

Ogdin, Carol Anne. Microcomputer Management and Programming. Englewood Cliffs, NJ: Prentice-Hall, Inc., 1980.

National Semiconductor Corporation. Linear Applications Handbook 2. Santa Clara, CA: National Semiconductor Corporation, 1976.

Texas Instruments Learning Center. Understanding Solid-State Electronics. Dallas, TX: Texas Instruments, Inc., 1978.

TEST

1. Why is higher resolution desirable in d-a or a-d converters?

2. What distinguishes a resistor-ladder network from other resistor networks that accomplish d-a conversion?

3. How does an op-amp transistor differ from a switching transistor?

4. What device is common to all ADCs and what is its function?

5. Define quantizing error. How is it different from linearity?

6. Describe the two types of counter-comparator ADCs introduced in the module. Which is superior?

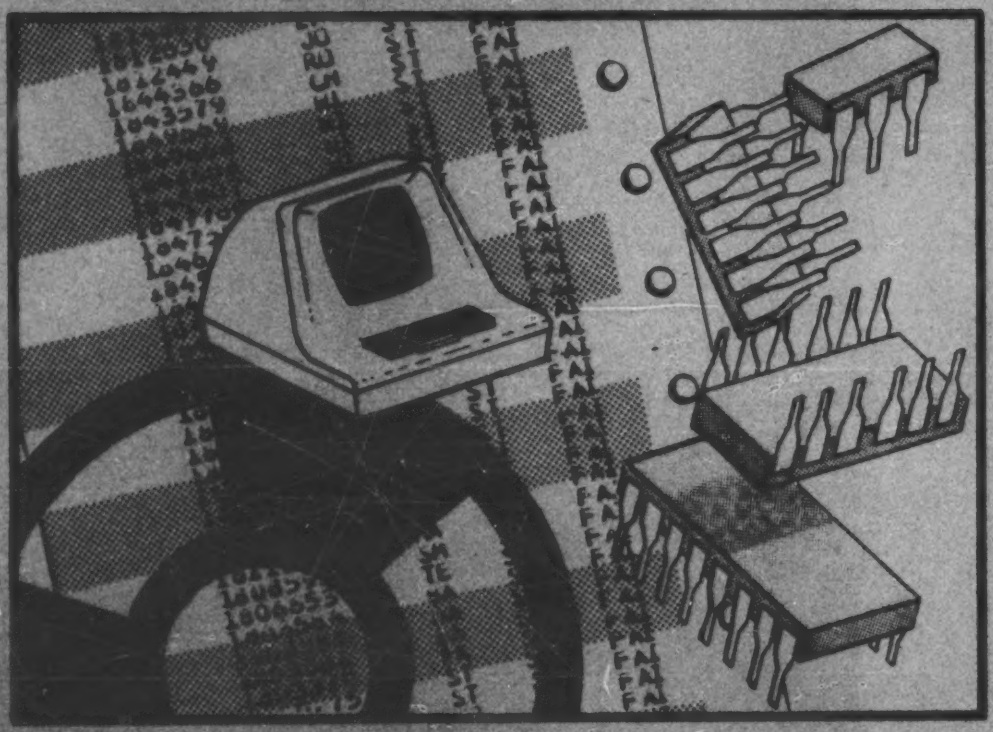
7. How does the successive approximation method of a-d conversion generate reference voltages?

8. What is the main disadvantage of simultaneous ADCs?

ENERGY TECHNOLOGY

CONSERVATION AND USE

MICROCOMPUTER HARDWARE



MODULE MH-05

DATA COMMUNICATION

TECHNICAL EDUCATION RESEARCH CENTER - SOUTHWEST
4800 LAKEWOOD DRIVE, SUITE 5
WACO, TEXAS 76710

INTRODUCTION

The various components of a computer must communicate with each other in order to perform a given task. Increasingly, applications require that the input/output components – or even mass storage or memory devices – be separated from the CPU by some distance. Examples of these applications include bank terminals that are connected to a central computer, telephone systems, instruments on Mars that transmit data to Earth, and implanted sensors within living organisms that communicate with laboratory monitor units. Special attention must be paid to the communication of data between various components.

The current trend toward distributed data processing also requires data communication. A local microcomputer may have all the necessary components on-site to input, process, and output data by connecting subsystems together in a network of intelligent terminals. Each local system may also have access to vast resources of additional data. Within a large building complex, for example, a microcomputer might have a dedicated function of controlling lighting, heating, and cooling in an energy-efficient manner. Periodically, this microcomputer subsystem might send routine data it has gathered during the controlling function to a central computer. The central computer then prepares financial reports and analyses for building management.

In this module, the student will be introduced to communications devices, techniques and data formats, as well as the jargon used when discussing data communication.

PREREQUISITES

The student should have completed Modules MH-01 through MH-04 of Microcomputer Hardware.

OBJECTIVES

Upon completion of this module, the student should be able to:

1. Define basic communication concepts such as medium, message, noise, channel, time-division multiplexing, and frequency-division multiplexing.
2. Distinguish between synchronous and asynchronous transmission.
3. Distinguish between baud and bits-per-second rates to describe channel capacity.
4. Define the following terms:
   a. Protocol.
   b. Modulation.
   c. Demodulation.
   d. Bandwidth.
   e. Signal-to-noise ratio.
   f. Attenuation.
   g. Amplitude modulation.
   h. Frequency modulation.
   i. Pulse modulation.
   j. Pulse-amplitude modulation.
   k. Pulse-width or pulse-duration modulation.
   l. Pulse-position modulation.
   m. Pulse-code modulation.
   n. Nyquist rate.
   o. Modem.

DATA COMMUNICATION

The primary function of a data communication system is to transfer information from one point to another by means of signals carried over a medium.

The medium (plural form is media) might be a wire over which electrical signals transmit information. Other media include radio waves and light waves that are guided by fiber-optic cable.

Students should notice that the term "medium" is used to describe both physical carriers, like wire, and electromagnetic phenomenon, like radio and light waves. Light can act as a medium through space with no guiding apparatus. For example, ships at sea can communicate by flashing lights in code. Light waves may also be guided through fiber optic cable – in which case, both the cable and the light become a medium over which information is communicated.

A distinction must be made between the medium and the intelligence signal that communicates the information (or message) that was intended. A signal does not always communicate information. Therefore, the term "signal" is not equivalent to the term "message."

Intelligent signals usually communicate a message in an orderly pattern or code. The signals are changes in the communication medium, usually either in the ...

a. strength, intensity, or amplitude of the signal;
b. by the frequency of the signal (its rate of variance through time); or,
c. by the phase of the signal (normally, a comparison between the variations in two signals).

Since changes in the communication medium may be caused by random, unwanted sources – like atmospheric conditions or impurities in the transmission medium – signal interruption may interfere with message clarity. Random or unwanted signals are called noise.

The process of encoding a message onto a medium is called modulation. The process of decoding the message is called demodulation.

CHANNELS AND BANDWIDTHS

The rate at which a medium can change often determines the volume of information it can communicate. This is why, for example, higher frequencies (faster rates of change) are required for television than for AM radio. This is because the television signal is more complex and is easier to communicate over a higher frequency. A typical AM radio signal varies 10,000 times a second, while a typical television signals varies 6,000,000 times a second – a rate 600 times greater.

The terms "channel" and "bandwidth" are used to define a specific portion of an information-bearing medium.

For example, frequencies between 54,000,000 and 60,000,000 cycles per second define television channel 2 in the electromagnetic spectrum. The two frequencies identify the exact position of the channel in the medium (the electromagnetic spectrum). The difference between the two frequencies is the bandwidth of the channel.

Bandwidth indicates the volume of information that can be carried within the channel. Most channels are designed to

be filled by the volume (width) of the signal a channel is designated to carry – with a small amount of space on either side to help isolate information from neighboring channels.

For example, a television signal requires a 6 megahertz bandwidth to communicate the entire signal. This allows 4.5 megahertz for the video portion, 50 kilohertz for the audio portion, and a small space on either side for isolation. ("Hertz" is the international term meaning cycles per second.)

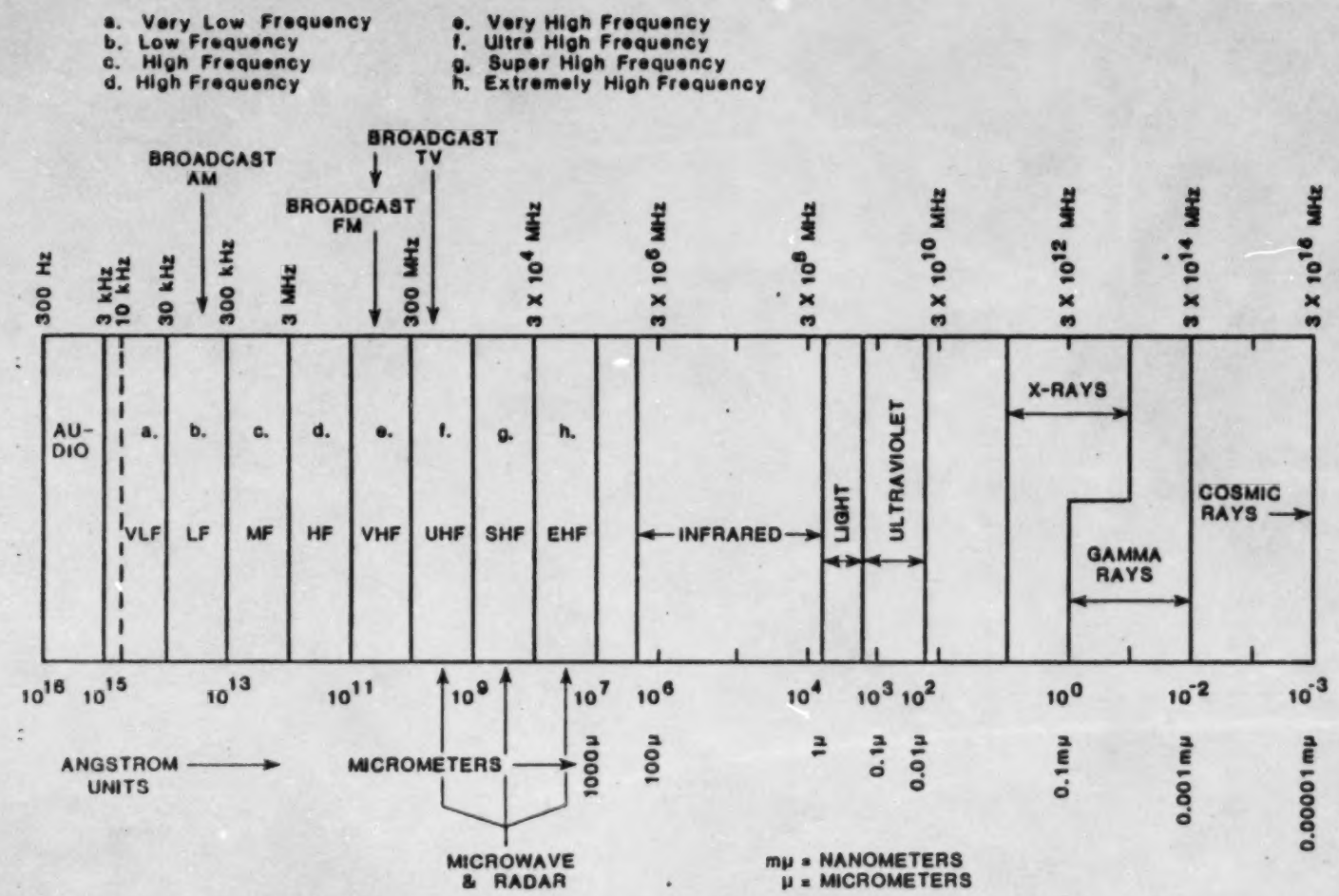
The Federal Communications Commission (FCC) governs the use of the electromagnetic spectrum in the United States in an effort to ensure that the needs of all interested parties are served.

Increasing demands for information (data) transmission in the United States have placed heavy expectations on this governing body. The FCC has the responsibility of keeping pace with rapidly-occurring technological developments.

Technically, the electromagnetic spectrum is infinite in both directions. Allocation of frequencies within the capabilities of current technology is very important, due to high demand for the airwaves. For this reason, techniques that result in more efficient use of available bandwidths are constantly being developed. These techniques include data compression – which reduces the volume of signals required to communicate a message – and multiplexing, which enables more than one message to be communicated simultaneously on the same channel.

Media (used for data communication between microcomputers and their peripherals) can be any portion of the electromagnetic spectrum, which is diagramed in Figure 1.

Basically, the only difference between the 60-cycle alternating current flowing through most homes in the United



148

149

Figure 1. A Representation of the Electromagnetic Spectrum.

States and radio, light, or X-rays, is the frequency at which the signals vary. All are manifestations of the electromagnetic spectrum; each segment of the total spectrum behaves differently and has peculiar characteristics. For this reason, certain parts of the total spectrum are more suitable for particular communication requirements than others.

For example, most of the spectrum has a tendency to travel through space, or radiate, as is the case with radio waves. Even household current generates a certain amount of radiation that can cause a hum (or interference) in a radio receiver placed near a wire carrying electricity. This extraneous noise can interfere with effective communication.

A narrow segment of the electromagnetic spectrum is visible as light. This segment is not affected by unwanted radiation and its rays will travel within the confines of a channeling device, such as glass fiber. The peculiar characteristics of this portion of the spectrum are one reason why fiber optics are becoming popular as a method of communicating data.

An example of the advantages of fiber optic cable can be found in standard telephone communication. Many tiny wires, each enclosed in a thin layer of insulation, are bound together in thick cables. Each wire carries a portion of a telephone conversation by electrical signals.

However, because of the radiation effect (mentioned earlier) and the close proximity of the wires, signals can become shared, resulting in crosstalk. Crosstalk can be heard as faint conversation in the background. Lightning and static electricity also contaminate the message signal, resulting in an audible crackling noise.

Light waves, however, are immune to these effects and can be sent through fiber optic cables in close proximities. Crosstalk and extraneous noise do not affect light waves.

The concept of noise is important to the field of communication. As shown in the previous example, the choice of medium can help to diminish the effect of noise. Often, the medium can be improved to shield or insulate it from the causes of noise. The message itself can help prevent unwanted noise – because of the characteristics of the actual signals and the encoding method.

One of the most important measurements of a communication medium is its signal-to-noise ratio. The higher this ratio, the less noise will be present at the end of the communication link.

The signal-to-noise ratio, or S/N ratio, is the relationship of the magnitude of the desired signal compared to the magnitude of the unwanted noise signal. The standard unit of measurement for S/N ratios is the decibel, abbreviated dB.

Figure 2 diagrams basic components of a communication system and indicates the points at which noise can contaminate the sytem.

Signal drain of the transmission medium upon the relative signal strength of the message is called attenuation. By using different types of modulation techniques or amplifiers, the effects of attenuation can be minimized.

As mentioned in the earlier module on analog/digital conversion, digital electronics have rapidly replaced analog systems in the world of computing. This is rapidly becoming the case in the world of communication also, although both methods will continue to be used for some while.

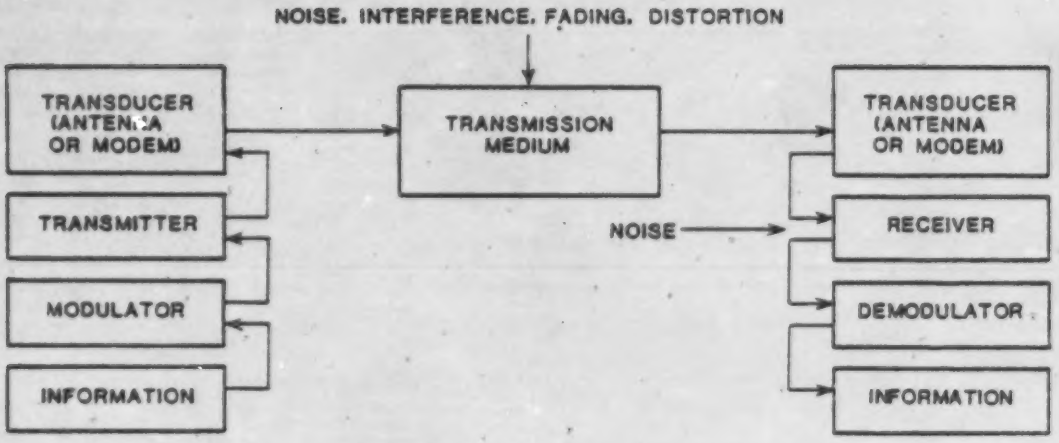


Figure 2. A Communication System Block Outline.

Hand-held digital microcomputer terminals have recently been introduced that communicate with a larger fixed computer by means of portable walkie-talkies. This exemplifies digital technology that is converted to analog AM or FM radio communication and then converted back to digital. These devices might be useful to technicians involved in energy conservation. The technicians might use walkie-talkies as a means of accessing and transmitting large amounts of data from outside locations to a fixed computer.

AM and FM stand for amplitude modulation and frequency modulation, two popular ways in which data are transmitted and received. As the name indicates, AM varies the amplitude, or strength, of a signal to communicate information. FM varies the frequency of a constant amplitude signal to communicate. These differences are diagramed in Figure 3.

Both methods are analog, or continuous wave, forms of communication. When these methods are used for data transmission over long distances – whether as radio waves or through coaxial cables – the strength of the signal

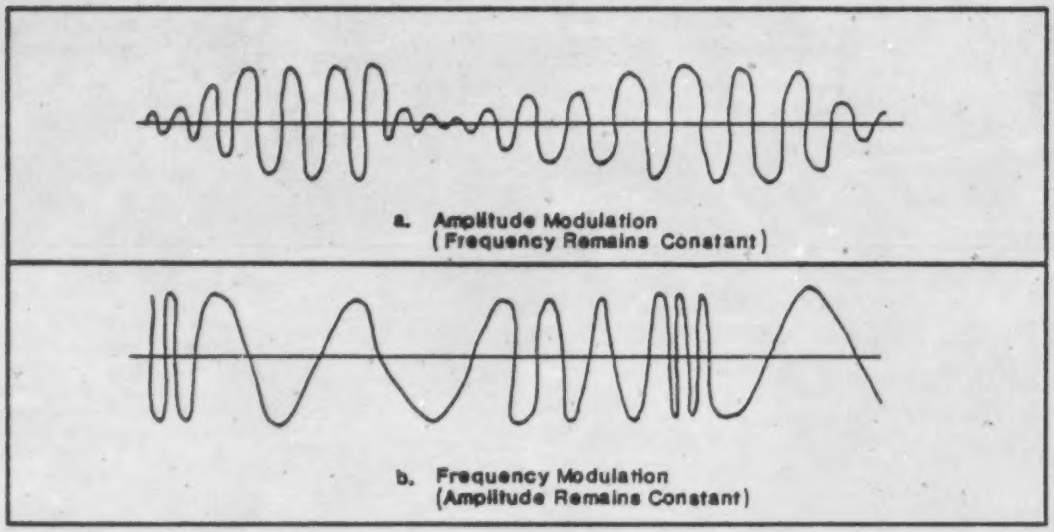


Figure 3. AM and FM Modulation Techniques.

deteriorates. Therefore, these signals must be amplified at various intervals to maintain the message.

Unfortunately, any noise that has entered the system will be amplified unless the signal is processed to improve its S/N ratio. If the noise level becomes too high, the information is lost.

Various other methods can be used to modulate information onto a medium.

For example, it can be mathematically proven that when a continuous, analog wave is sampled at a rate equal to twice its highest frequency component, the wave can be reconstructed at the receiving end of a transmission line by communicating only the sample information.

This has led to the development of communication techniques that, although still analog, represent an improvement over continuous wave communication because they facilitate multiplexing and data compression.

Figure 4 diagrams three methods of pulse modulation (PM) and an equivalent continuous wave (CW).

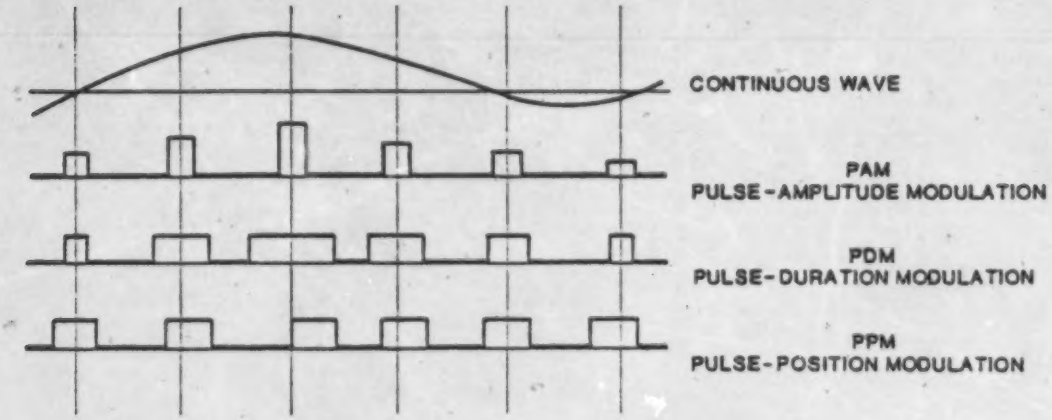


Figure 4. Diagram of Three Pulse Modulation (PM) Techniques.

In pulse-amplitude modulation (PAM), the amplitude of the pulse represents the sample amplitude of the continuous wave at a given time. This information enables the receiving device to interpolate missing information to reconstruct the entire signal. This is similar to plotting a curve by knowing only a minimum amount of coordinates. However, the sampling rate must be at the specified level to ensure the accuracy of the recreation.

By placing a different message in the holes between the pulses, two messages may be combined, or multiplexed, within the same channel. This method is also known as frequency division multiplexing, or FDM.

There are two other methods of pulse modulation. One is pulse-duration modulation (PDM), (also called pulse-width modulation [PWM]), which varies the width (length in time) of the pulses to correspond to the continuous wave. The other method is pulse-position modulation (PPM), which varies the position in time of the pulse to communicate the information.

When multiple messages are communicated on the same channel using these techniques, the process is called time division multiplexing (TDM).

In all of the preceding methods of modulation, analog techniques were used. The advent of digital communication in the form of pulse-code modulation (PCM) has resulted in high rates of communication with extremely high S/N ratios.

The two terms used in data communication to indicate the rate or speed of data transfer are "baud" and "bits-per-second." A baud equals one-half of a full cycle; therefore, the baud rate equals two times the bandwidth of the signal. This rate, also called the Nyquist rate, is hard to achieve because the equipment must accurately detect changes every half-cycle. The precision equipment required to achieve the Nyquist rate is expensive.

In digital communication, the actual rate of transfer is normally defined in bits-per-second, or bps. The information to be communicated has been converted to binary code. Each pulse represents one binary condition, while the absence of a pulse indicates the other binary condition. By 1976, bps rates of 16.4 billion bits-per-second had been achieved.

In pulse-code modulation, the equipment must only detect the presence or absence of pulses. Noise that enters the system may deteriorate the level of the ON pulses, and add a small amount of signal where there should be none (in the OFF

segments). However, it is difficult to degrade the integrity of a PCM signal to the extent that noise affects the intended recognition. Thus, the PCM S/N ratio is very high.

In addition, each of the amplifying stations along the route of long-distance PCM lines generates a new perfect signal, restoring the level of ON pulses and eliminating signal from the OFF segments. The result is very effective long-distance communication through PCM techniques.

PCM is also adaptable to data compression and multiplexing techniques.

As mentioned in earlier modules, data transmission may be either serial or parallel, synchronous or asynchronous. Most communication links are serial due to the high cost of multiple lines required to conduct parallel transmission. Normally, parallel communication is only cost-effective when transmitting over short distances, such as connecting computers within the same building.

Asynchronous formats may vary over an arbitrary time interval between signals. For example, in digital communication, the code representing an ASKII character might be sent with special characters inserted before and after the character to represent the beginning and end of the character. The time that might lapse before the next signal is transmitted or received may vary in asynchronous communication.

Synchronous communication provides more efficient channel utilization. However, synchronous communication is more adaptable for block-oriented applications where large amounts of data are sent in blocks. Each group is identified by beginning and ending data. In data networks, these groups are of fixed lengths and are called packets.

The characters within these blocks are transmitted in succession, with no individual beginning or ending definitions. Characters are transmitted at a specified rate that is normally governed by a clock.

Any pause between characters causes receiving equipment to lose synchronization for the remainder of the communication.

The trend in worldwide data networks is toward synchronous communication that uses a common, international clock system that is based on a world time standard.

Data networks are rapidly being built to take advantage of the merging fields of voice, image, data, and message communication as microcomputers and digital electronics become used in more and more communication applications.

Satellite transmission techniques have improved rapidly. Advances such as the space shuttle, which will enable more economic satellite placement and maintenance, should help increase the importance of satellite transmission techniques.

Meanwhile, land-based systems founded on fiber optic links are rapidly replacing cable and microwave applications.

Heavily-multiplexed networks require standardized rules of interchange, called network protocols, to identify the nature and destination of specific information.

The term "simplex" has been used to identify a channel in a data network that allows information transfer in only one direction. (It has also been used to distinguish a line that has not been multiplexed.)

A half-duplex line allows communication in either direction – but in only one way at a time.

A full-duplex line allows simultaneous communication in both directions.

At either end of the line, a modem – short for modulator/demodulator – might be required as an interface to the computer.

As formats become standardized and microcomputers are integrated into every level of communication, fewer modems will be required as translaters.

Digital switching techniques have also improved channel utilization by pinpointing and using every available position within a transmission medium.

For example, in a normal phone conversation using full-duplex lines, the transmission portion of a person's equipment is not being utilized when he is listening, and vice versa.

Large corporations can save money by allowing a digital computer to control the utilization of its available communication channels and, therefore, do more with less.

This is only one example of many methods that will become increasingly commonplace as the era of digital communications arrives.

EXERCISES

1. What ratio is measured in decibels? What is the abbreviation for decibel?
2. What do the following abbreviations represent?
   a. FCC c. PCM e. TDM g. PPM i. AM
   b. PAM d. PDM f. CW h. FM j. FDM
3. Define the following terms:
   a. Crosstalk.
   b. Multiplexing.
   c. Modulation.
   d. Protocol.
4. (from Figure 1):
   a. In what frequency ranges do most of today's communication transmitters and receivers operate?
   b. The electromagnetic spectrum has been divided up into bands that are _____ hertz wide.
5. At what points does noise generally enter a communication system?
6. A baud equals _____ cycle.
7. The Nyquist rate equals _____ the bandwidth.
8. What is the difference between bandwidth and channel?

LABORATORY MATERIALS

Microcomputer (Commodore-KIM-1).
Power Supplies (+5, +12, +9, or 9-volt battery).
Cassette tape recorder.
Oscilloscope.
Small strand of fiber optic cable.

The following transistors:

1	2N2907	1	2N2647 or 2N4891
1	FPT-100 photo transistor		

The following integrated circuits:

1	MC1458	1	LM386

The following resistors:

1	1.2 k ohm	1	5 k ohm	1	100 ohm
1	2.4 k ohm	1	15 k ohm	1	100 k ohm
1	3.2 k ohm	1	200 ohm	1	10 m ohm
2	10.0 k ohm	1	1 k ohm		

The following capacitors:

1	0.047 μF	4	0.1 μF	1	220 μF
1	0.022 μF	1	10.0 μF		

1 8-ohm speaker

1 GaA (900 nm) LED

LABORATORY PROCEDURES

In this laboratory, the student builds a fiber optic communication link, and uses the KIM-1 microprocessor to generate data. These data are then checked with an oscilloscope.

Briefly, the advantages of fiber optic cable are the following:

1. Glass and plastic fibers are smaller and lighter than their copper counterparts, yet have a higher bandwidth.
2. Eventually, the price of fiber optic cables will be much less than the price of metal conductors. This is because the basic raw material for fiber optic cable is sand.

3. Fiber systems are impossible to jam, and the signals are very difficult to intercept.
4. Because of the non-electrical nature of light-frequency communication, fiber optic systems are immune to electromagnetic interference. Fiber optic systems can be used in environmentally dangerous situations, such as explosives dumps or gas-filled rooms, with no danger of electrical spark. They are also useful in high-energy-producing areas, such as electric generating plants, where electromagnetic interference is high.
5. Glass fiber research is continually developing fibers with lower and lower attenuation. Attenuation in glass fibers is now well below 1 dB/kilometer, making possible long-distance links with fewer repeaters.
6. Glass fibers are stonger than steel wire of the same diameter.

These are some of the reasons why fiber optic systems are now replacing wire and microwave links.

Most communications links require graded-index fibers, which have light transmission characteristics that enable high data rates. However, for the small transmission distance required for this lab, any grade of fiber optic cable should work.

Some of the current manufacturers or suppliers of fiber optic cable include Edmund Scientific, ITT, Corning, Valtec, Siecor, DuPont, and Quartz.

Fibers may be cut with a knife or razor blade. If possible, examine the cut with a 50-power, phono-stylus microscope (or a printer's eyepiece) to make sure the cut has no burrs or nicks to diffuse light flow. Lengths of one or two feet will be sufficient for this laboratory.

An easy way to attach the fiber to the LED is by heating an awl and melting a small hole (about the same diameter as the fiber) through the epoxy protection directly over the LED chip. Be sure to test the chip after producing the hole to ensure a bright light source. The fiber may be secured in the epoxy with Eastman 910 Adhesive, or an equivalent.

When attaching the fiber to the phototransistor, make sure that the transistor is shielded from any spurious light.

The schematic diagram for the transmitter and receiver circuits is shown in Figure 5.

The 0.1 μF capacitor across the power supply pins of the op amp in the receiver helps prevent violent oscillations. Never use an earphone to test the output of an untested receiver, as the sound pressure level might exceed one's pain threshold.

This transmitter communicates by PFM, or pulse frequency modulation. All bursts of light that indicate information should have the same amplitude. This allows a threshold circuit in the receiver to block lower amplitude noise.

When unmodulated, the transmitter sends a constant series of pulses at the same frequency. Signals applied to the transmitter caused proportional changes in frequency to communicate information.

Changing R4 adjusts the comparator; R2, the gain; and R5 and C2, the tone response.

Use the sound synthesizer program from the laboratory from Module MO-01, "Computer Codes," to generate signals at ports listed in the Data Table, and then alter the addresses as described to generate different op codes. Describe your results in the Data Table.

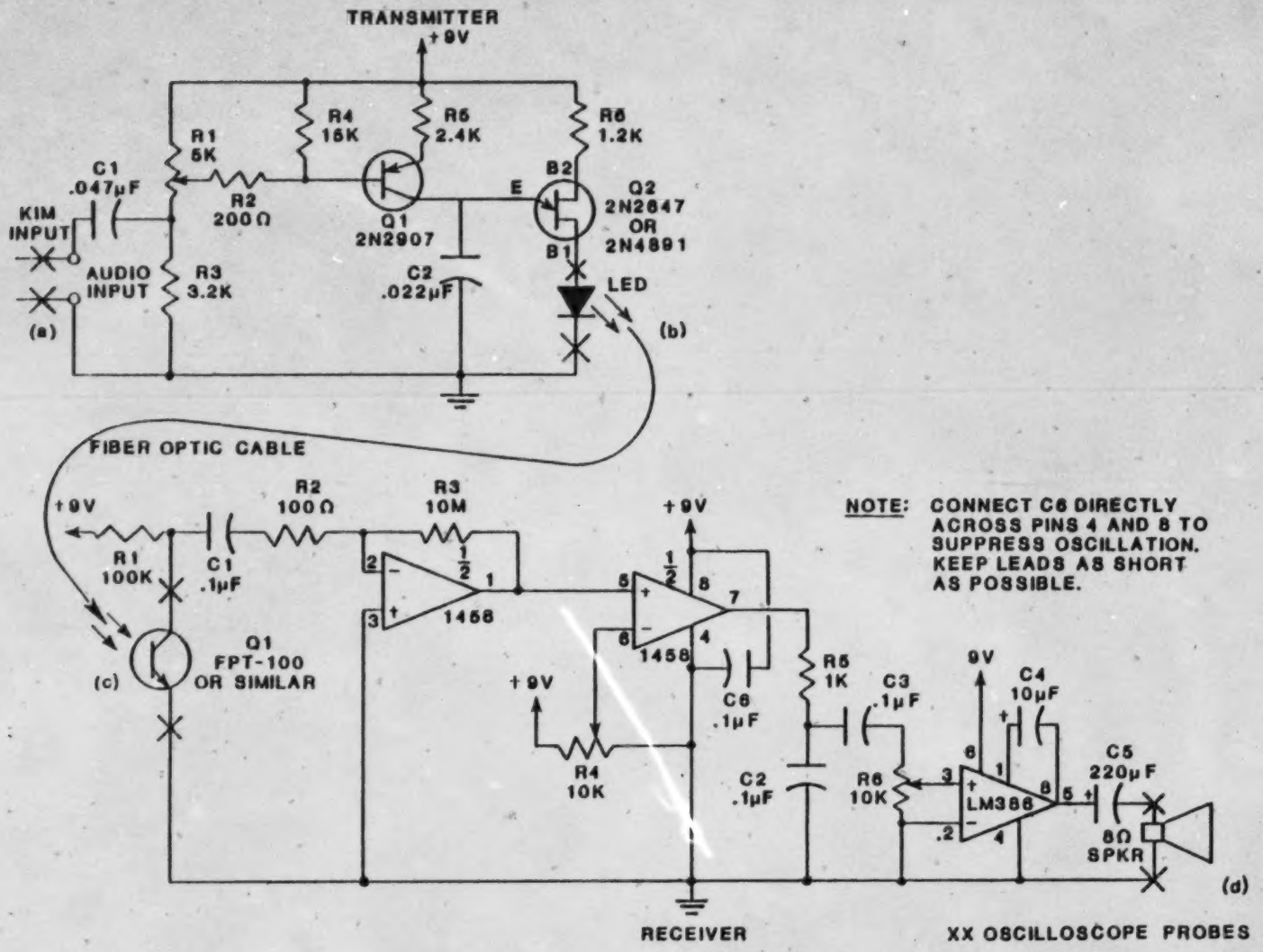


Figure 5. PFM Fiber Optic Transmitter/Receiver System.

Use the oscilloscope to look at ...

a. the signal being generated by KIM.

b. the pulse signal to the LED.

c. the output from the phototransistor.

d. the output to the speaker.

Use the following ranges on the oscilloscope at each point:

a.

b.

c.

d.

DATA TABLE

PA line	A.	B.	C.	D.
0				
1				
2				
3				
4				
5				
6				
7				

Address	A.	B.	C.	D.
0000				
0001				
0002				
0003				
0004				
0005				

REFERENCES

Davenport, William P. Modern Data Communication: Concepts, Language, and Media. New York: Hayden Book Company, Inc., 1971.

Davies, D.W., et al. Computer Networks and Their Protocols. Chichester: John Wiley & Sons, Inc., 1979.

Doll, Dixon R. Data Communications: Facilities, Networks, and Systems Design. New York: John Wiley & Sons, Inc., 1978.

Freeman, Roger L. Telecommunication Transmission Handbook. New York: John Wiley & Sons, Inc., 1975.

Housley, Trevor. Data Communications and Teleprocessing Systems. Englewood Cliffs, NJ: Prentice-Hall, Inc., 1979.

Kuo, Franklin F., ed. Computer Communications Networks. Englewood Cliffs, NJ: Prentice-Hall, Inc., 1973.

Miller, Gary M. Handbook of Electronic Communication. Englewood Cliffs, NJ: Prentice-Hall, Inc., 1979.

Mims, Forrest M. "Optical Fiber Communications, Parts I and II." Popular Electronics. May and June, 1980.

Peebles, Peyton Z. Communication System Principles. Reading, MA: Addison-Wesley Publishing Co., 1976.

Roden, Martin S. Analog and Digital Communication Systems. Englewood Cliffs, NJ: Prentice-Hall, Inc., 1979.

Sippl, Charles J. Data Communications Dictionary. New York: Van Nostrand Reinhold Company, 1976.

TEST

1. What inherent characteristics enable digital communication to have a lower signal-to-noise ratio than analog communication?

2. Why is it important to utilize communication channels efficiently? Name and explain two methods of doing so.

3. Why is attenuation undesirable?

4. Identify the modulation techniques below as being analog or digital, and explain why.

a. Pulse-Amplitude Modulation

b. Pulse-Duration Modulation

c. Pulse-Position Modulation

d. Amplitude Modulation

e. Frequency Modulation

f. Pulse-Code Modulation

5. Why is serial transmission used more frequently than parallel transmission?

6. Define the terms "simples," "half-duplex," and "full-duplex."

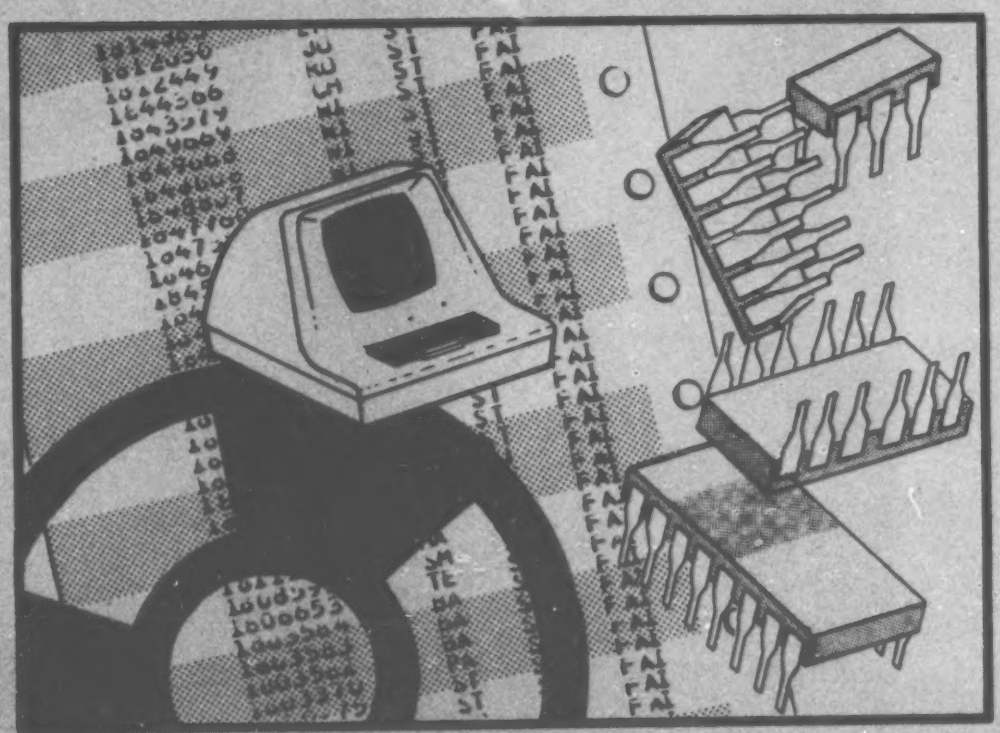
7. What does modem represent and what is its purpose in a communication system?

8. What is the difference between a medium and a message?

ENERGY TECHNOLOGY

CONSERVATION AND USE

MICROCOMPUTER HARDWARE



MODULE MH-06

BUS SYSTEMS

TECHNICAL EDUCATION RESEARCH CENTER - SOUTHWEST
4800 LAKEWOOD DRIVE, SUITE 5
WACO, TEXAS 76710

INTRODUCTION

The bus is the "spinal cord" of a computer. It provides the main line of communication between intelligence in the system (the CPU) and all other parts of the computer.

The characteristics of each microcomputer's bus configuration are important. These characteristics determine the ease with which a system can be expanded to include more memory, pheripherals, and accessories now available from many different manufacturers.

In this module, the student will take a closer look at the bus system to see what signals travel on it and how individual components are connected to it, both logically and electrically.

PREREQUISITES

The student should have completed Modules MH-01 through MH-05 of the course Microcomputer Hardware.

OBJECTIVES

Upon completion of this module, the student should be able to:

1. Define and illustrate the term "bus."
2. Cite and describe the functions of the three major parts of a computer bus.
3. Identify the following terms:
   a. Motherboard
   b. Plug-compatible

c. Bus drivers
d. LSI-11 bus
e. S-100 bus
f. Multibus
g. STD bus

CHARACTERISTICS OF A BUS

The bus system of a microcomputer is the physical path over which the electrical signals carrying information flow between the various components of the computer.

It is important to know the characteristics of the bus structure for several reasons. One reason is that, when purchasing a microcomputer or microcomputer-based system, the specific configuration of the bus system may play an important role in the future expansion of the system.

Many manufacturers have adopted common busing system configurations (such as the S-100 bus, which will be discussed later). As a result, there are many more readily available — and less expensive — components for microcomputers using these busing systems than those that utilize other, less common systems. These modular, standardized design schemes for buses used by most manufacturers ease the problems associated with interfacing computer peripherals* to a specific microprocessor.

SIGNALS

When repairing or troubleshooting microcomputers it is necessary to know the details of the internal communication system of the computer.

The specific signals communicated within the computer fall into three main categories: address signals, data signals, and control signals. As a result, three separate

*Peripherals: Devices connected to a computer to provide communication (as input and output) or auxiliary functions (as additional storage).

parallel "pipeline" systems exist within the microcomputer to carry these signals back and forth.

Address Lines

Address lines carry the signals that convey the location of information in RAM or ROM memory, either internally or as a peripheral memory device. The information contained within the memory at these locations is required for processing, or they give direction to the computer. Most microcomputers have either 8 or 16 separate lines in the address bus.

Data Lines

Data lines — which normally total 8 — carry the instructions and/or data that can be found at the addresses which have been selected by the signals on the address bus. Depending on the structure of the microprocessor (which controls the flow of data on the data bus), these lines can be bidirectional and dedicated to a specific function, such as MEMORY-DATA-IN only, or I/O-DATA-OUT only.

Control Lines

The control lines include the master clock line, which coordinates the operations of the various components within the microcomputer. They also include other dedicated control lines for functions such as READ, WRITE, HALT, RESET, INTERRUPT REQUEST, and FLAGS.

MOTHERBOARD

Most microcomputers are built around a "motherboard." A motherboard consists of a printed circuit board that has many connectors for other printed circuit boards. These circuit boards also have a common busing scheme (including identical operating voltages) which might contain the ICs, capacitors, resistors, and so forth, for a real-time clock, additional memory, speech recognition, peripheral control circuitry, or many other functions.

The KIM-1 is a single-board computer, although an expansion motherboard is available for the KIM series.

The term "plug-compatible" is commonly used to identify boards that will function properly without modification when connected to specified systems.

COMPARISON OF BUSING SYSTEMS

The S-100 bus, the first widely accepted busing standard, was originally developed by MITS, Inc. for the 8080 microprocessor-based Altair computer. The name is derived from the fact that there are 100 pins in the connectors for S-100-compatible boards. Although other busing systems are gaining in popularity, S-100-based systems and accessories are widely available.

Table 1 shows a comparison between the S-100 and three other common bus systems, which will be discussed later.

TABLE 1. COMPARISON OF POPULAR BUS SYSTEMS.				
Characteristics	S-100	STD	LSI-11	Multibus™
Signal Distribution	100	56	72	86
Address	16	16	18	20
Data	8+8	8	—	16
Control	38	21	18	18
Interrupts	8	1	1	8
Grounds	2	4	8	8
Power	4	6	9	16
Spares	16	0	18	0
Features				
#Voltages	3	3	3	5
Regulation	No	Yes	Yes	Yes
Battery Lines	0	0	2	(Aux)
Bus Arbitration	No	No	No	Yes
Dimensions				
Height, in	5.3	6.5	8.5	6.75
Width, in	10	4.5	5.2n	12

TIMING AND SYNCHRONIZATION

The terms "enable" and "disable" are helpful when analyzing the way a bus system operates. Many times, the control signals on the various control lines allow functions to occur only when they are enabled by a voltage level going HIGH or LOW.

In the case of bidirectional lines in a data bus, the control lines "READ" and "WRITE" (for example) indicate what the intended purpose is to be for the signals present on those data lines.

When the READ line is enabled, the signals on the data bus represent information going to the CPU from a memory location. If the same data were on the bus when the WRITE line was enabled, the signals would be interpreted as intended for storage into memory at the address currently being indicated by the address bus.

The timing and synchronization of these operations are critical. Since it is not possible to be reading and writing data on the data bus at the same time, the READ and WRITE functions normally share the same control line, with one of the functions operating under reverse logic from the other.

Figure 1 illustrates this principle. If the READ function is enabled when the control line is HIGH, the WRITE function is enabled when the control line is LOW.

Figure 2. The Status of One Control Line Indicates the Destination of Data on the Data Bus.

The cycling of the master clock line generally functions as a trigger that initiates the operations within a computer.

Buffers and latches serve as holding tanks for data so that data intended for the data bus as information going to to the CPU is held until the completion of a WRITE cycle, for example.

Another type of device, called a 3-state (or tri-state transistor) is also used in data bus structures. These devices can operate at normal high/low logic levels, but can also operate in a high-impedance state that is interpreted as NO DATA by any device trying to read data from the line when it is in that state. This ability serves as an added method of coordinating the delicate timing and synchronization of signals on common buses.

Of course, individual lines could be run for each necessary function, eliminating the need for the common pathways that serve as bus systems. However, each square micron of space of an integrated circuit chip is valuable, and each area devoted to bus circuits takes away space for more logic circuitry.

Because of the small distances required for the signals on an ON-chip bus to travel, the signals themselves do not need to be very powerful. In designing microcomputer systems, engineers have kept most of the ICs that have to communicate with the CPU in close proximity to the microprocessor chip. As a result, many of the signals eminating from the 40 pins on a microprocessor chip travel back and forth between immediately adjacent ICs and nowhere else.

Short distances keep signal power requirements lower and improve operational speed. However, in some applications, especially I/O, the strength of the signals on the bus must be increased by devices known as "bus drivers."

Most general purpose microcomputers are equipped with bus drivers powerful enough to successfully communicate with as many devices (receivers) as necessary in a normal application.

BUS INTERFACING

Prior to the purchase of peripheral equipment and to an attempt at interfacing, one should always check the following specifications and considerations:

- The characteristics of the signals.
- The timing relationships of the signals.

In addition, the physical means of interconnection, (bus connectors) as well as the physical dimensions of the modular units, must be compatible with the system in question.

Bus Connectors

Bus connectors can range from 50 to 100 pins, containing not only the address, data, and control lines but also power connections and pins for future expansion.

Many systems have helped diminish the adverse effects of obsolescence by allowing such "spare" pins. For example, a majority of microprocessors are based around an 9-line data bus. Busing systems which have the pins for a 16-line data bus make upgrading as easy as pulling an old printed circuit board and replacing it with the updated version.

Physical Dimensions

The physical dimensions of the board itself are defined in terms of the height of the card (or the distance between the connector edge and the one opposite it) and the width (or the distance between the two remaining edges). Almost all computer printed circuit boards have all the connectors on one edge.

Types of Interfacing Bus Systems

Digital Equipment Corporation, manufacturer of the famous PDP series of minicomputers, has chosen the LSI-11 bus system for its microcomputer line. The LSI-11 evolved from the company's experience with the earlier UNIBUS, although the two are not compatible. The LSI-11 bus is TTL compatible.

The MULTIBUS system was developed by Intel in 1973. It is also TTL compatible. The MULTIBUS system is unique in that it contains a 60-pin auxiliary connector in addition to its 86-pin standard connector. The auxiliary connector can be used for custom designs and for automatic testing.

The STD bus, developed by the Pro-Log Company, and the S-100 mentioned earlier are both designed primarily for 8-bit microprocessors (microprocessors that use 8-bit words). For this reason, some experts feel that the MULTIBUS and LSI-11 will become the dominant busing systems in the future.

Special interface cards are being developed that will allow microcomputer users to interconnect peripheral cards with uncommon bus systems. It is likely that the most popular bus interface card will be an S-100-to-MULTIBUS interface

because of the wide range of S-100-based cards in existence. However, since the S-100 bus is slower, it will be the limiting factor in the speed of such an interfaced system.

Documentation

Documentation is an important factor in interfacing bus systems. In addition to complete timing diagrams, block diagrams indicating the number of lines per bus (whether or not they are bidirectional) and the modular components they interconnect should be shown. Diagrams labeling all pins should be included.

In modern microcomputer systems, with many interconnected peripherals having features such as VECTORED INTERRUPTS and MEMORY-MAPPED I/O, the control of a bus at a given time might lie with the CPU, or it might lie with an external device.

However, the student should be aware of the variations in terminology in descriptions and diagrams.

For example, Figure 2 below illustrates how the three major bus systems within all microcomputers are linked together.

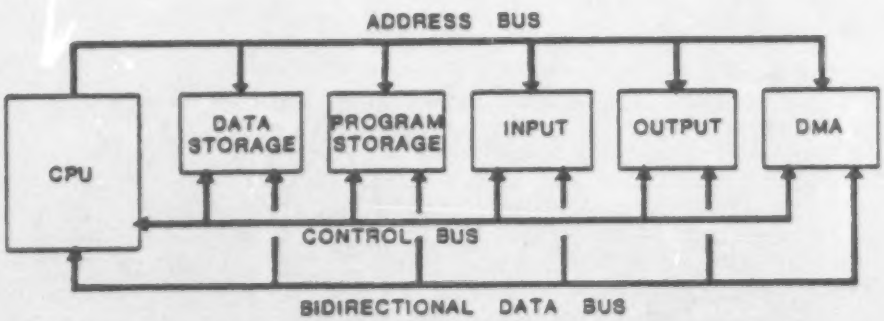
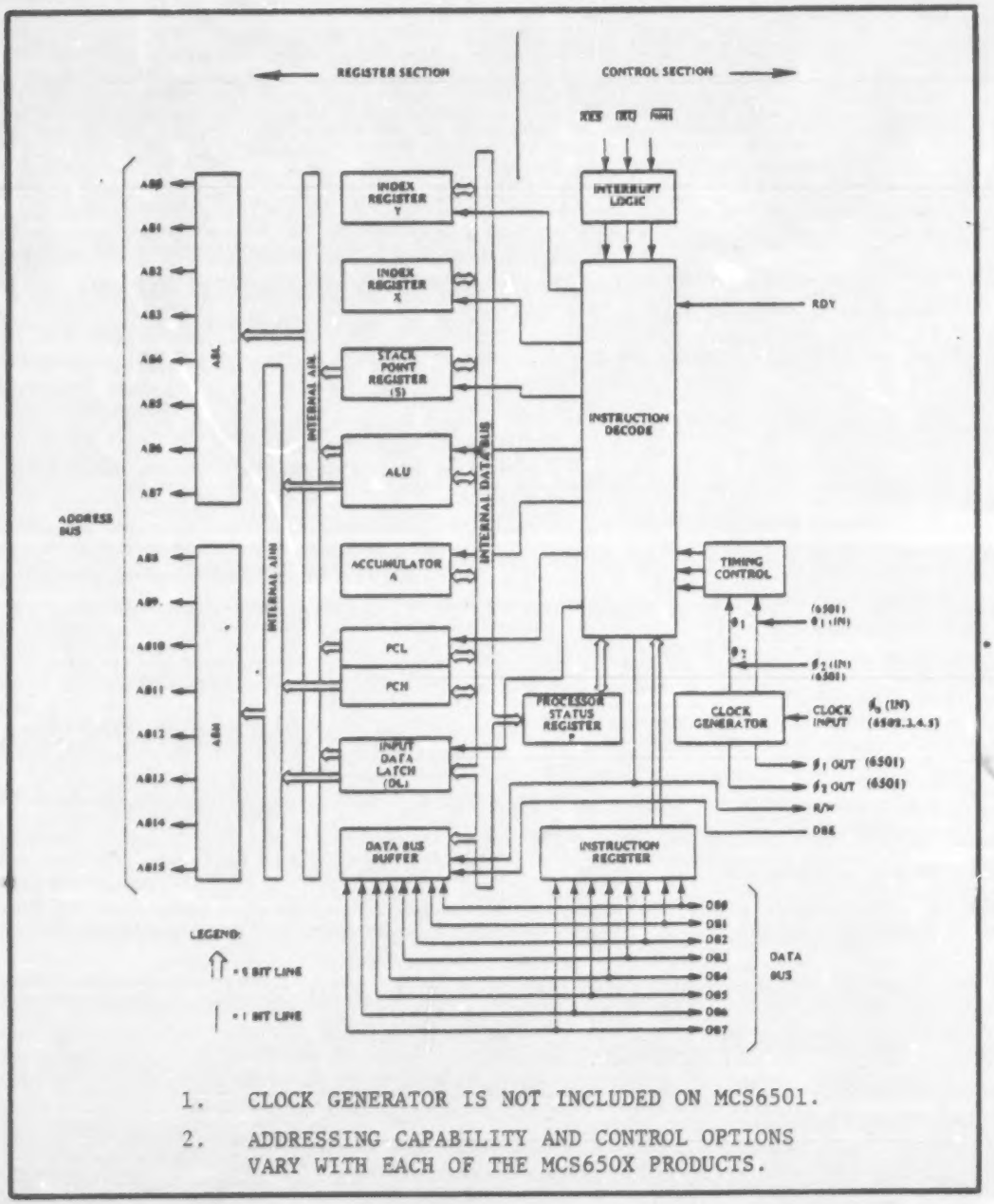


Figure 2. Relationship Between Three Major Buses and the Components of a Microcomputer. (DMA indicates peripherals with Direct Memory Access.)

Compare this figure with the system diagram for the KIM microprocessor in Figure 3. This.is typical of the way the internal parts of the CPU will be shown as individual function blocks, along with buffers, latches, and so forth.



1. CLOCK GENERATOR IS NOT INCLUDED ON MCS6501.
2. ADDRESSING CAPABILITY AND CONTROL OPTIONS VARY WITH EACH OF THE MCS650X PRODUCTS.

Figure 3. MCS650X Internal Architecture.

EXERCISES

Using Figure 3 of this module, answer the following questions for practice in reading diagrams.

1. How many address lines are shown on the address bus? _____ They are labeled _____ through _____.
2. How many lines are shown in the data bus? _____ They are labeled _____ through _____.
3. Which block appears to be the primary component of the CPU? The _____.
4. Which of the following component blocks shown in the diagram could also be part of the CPU?
   - _____ a. Program Counter Low (PCL)
   - _____ b. Program Counter High (PCH)
   - _____ c. Accumulator (A)
   - _____ d. Arithemtic Logic Unit (ALU)
   - _____ e. Stack point register
   - _____ f. Index registers (Y and Z)
   - _____ g. Interrupt logic
   - _____ h. Process status register
   - _____ i. Timing control
5. The reason the address bus is divided into sections is because _____.
   a. There are two sets in case one fails; there will be another address bus ready.
   b. This computer uses an 8-bit word, and by using two words to select a memory address – (Address Bus High (ABH) and Address Bus Low (ABL) – the computer can use a much larger memory than if it only had one 8-bit address bus.

LABORATORY MATERIALS

Microcomputer (Commodore KIM-1)
ASKII keyboard or supply board kit

LABORATORY PROCEDURES

Bus structures play an integral role in the expansion of microcomputer systems, as well as in the interfacing of multiple microcomputers.

As mentioned in the module, those systems having the widest distribution (especially in the hobbyist market), such as the S-100 bus, will have the broadest range of expansion and interfacing data available.

Many microcomputer magazines are available, featuring a wide variety of articles each month on a specific expansion/interface technique.

Since the KIM-1 was one of the first single-board microcomputers available, many hobbyists and computer enthusiasts took advantage of its low price and simple construction to become more familiar with microcomputers.

(For example, articles by Don Lancaster and Jim Trageser in the magazine Kilobaud show how to add an ASKII keyboard and video display to the basic KIM-1.)

PROCEDURE

The addition of the ASKII keyboard is desirable to improve the ease with which human operators "communicate" with the microcomputer. Figure 4 illustrates how the interconnection is made.

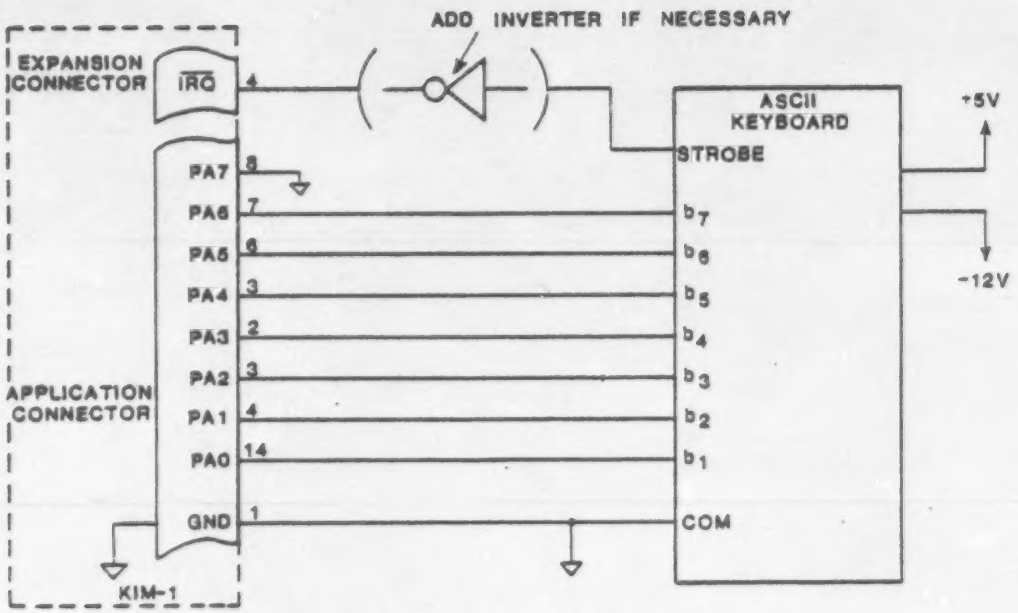


Figure 4. Diagram for the Addition of ASKII Keyboard to KIM.

The keyboard used in Figure 4 is a 7-bit parallel output keyboard based on the ASKII code described in an earlier module.

Connections are made using ports PA0 through PA6 on the application connector.

Notice that PA7 has been connected to ground. This connection, called "tieing a line low," is required to prevent the KIM from interpreting the signal on this line as a cursor bit.

The particular keyboard used should have positive, rather than negative, true logic. The strobe output from the keyboard should normally be in the HIGH STATE and PULSE LOW when the keyboard is active in order to work properly when connected to the $\overline{IRQ}$ (Interrupt Request Port) on the KIM expansion board. If it is not, it can be inverted.

The only additional connections required are the power circuits.

(Kits for ASKII keyboards or supply boards can usually be obtained for around $50.)

An article in the June 1977 Kilobaud illustrates the addition of a CRT to the KIM.

If the TVT unit described in this article is used, the strobe output from the keyboard in the diagram should be connected to TVT-6L pin 3, while the output pin 4 is connected to the KIM $\overline{IRQ}$.

In addition, the trace connecting pin 9 of IC-4-to-ground should be cut; and a jumper should be soldered between pin 8 and output pin 4, and between pin 9 on the IC and input pin 3.

REFERENCES

Bibbero, R.J. Microprocessors in Instruments and Control. New York: John Wiley and Sons, 1977.

Hall, Douglas V. Microprocessors and Digital Systems. New York: McGraw-Hill, 1980.

Lesea, A. and Saks, R. Microprocessor Interfacing Techniques. Berkeley, CA: Sybex, Inc., 1977.

MOS Technology, Inc. MCS6500 Microcomputer Family Hardware Manual. Norristown, PA: MOS Technology, Inc., 1976.

Ogdin, Carol Anne. Microcomputer Management and Programming. Englewood Cliffs, NJ: Prentice-Hall, Inc., 1980.

Peatman, John B. Microcomputer-Based Design. New York: McGraw-Hill, 1977.

Perone, S.P. and Jones, D.O. Digital Computers in Scientific Instrumentation. New York: McGraw-Hill Book Co., 1973.

Soucek, B. Microprocessors and Microcomputers. New York: John Wiley and Sons, 1976.

Toong, H.D. "Microprocessors," Scientific American, 237 (3), 146 (1977).

Zabro, R. Microprocessors. Berkeley, CA: Sybex, Inc., 1977.

TEST

1. Name the three types of internal buses in a microcomputer.
   a. _____
   b. _____
   c. _____

2. The device that features multiple connectors for interfacing peripheral cards with common busing schemes is called a _____.

3. These devices hold data until it is time for them to be placed on the data bus. _____.
   a. Multiplexers
   b. Buffers
   c. Drivers
   d. Tri-state

4. These devices amplify signals on the bus. _____.
   a. Tri-state
   b. Buffers
   c. Microns
   d. Drivers

5. These devices have a high-impedance state which corresponds to NO-data. _____.
   a. Multiplexor
   b. Driver
   c. Tri-state
   d. Microns

6. Check the following functions that were mentioned specifically in the module as being control functions:

____ a. READ
____ b. HALT
____ c. RESET
____ d. CLOCK
____ e. MOVE
____ f. INTERRUPT REQUEST
____ g. FLAGS
____ h. BUS
____ i. KEY

7. Fill in the blanks of the following questions with the letters which correspond to the proper bus system.

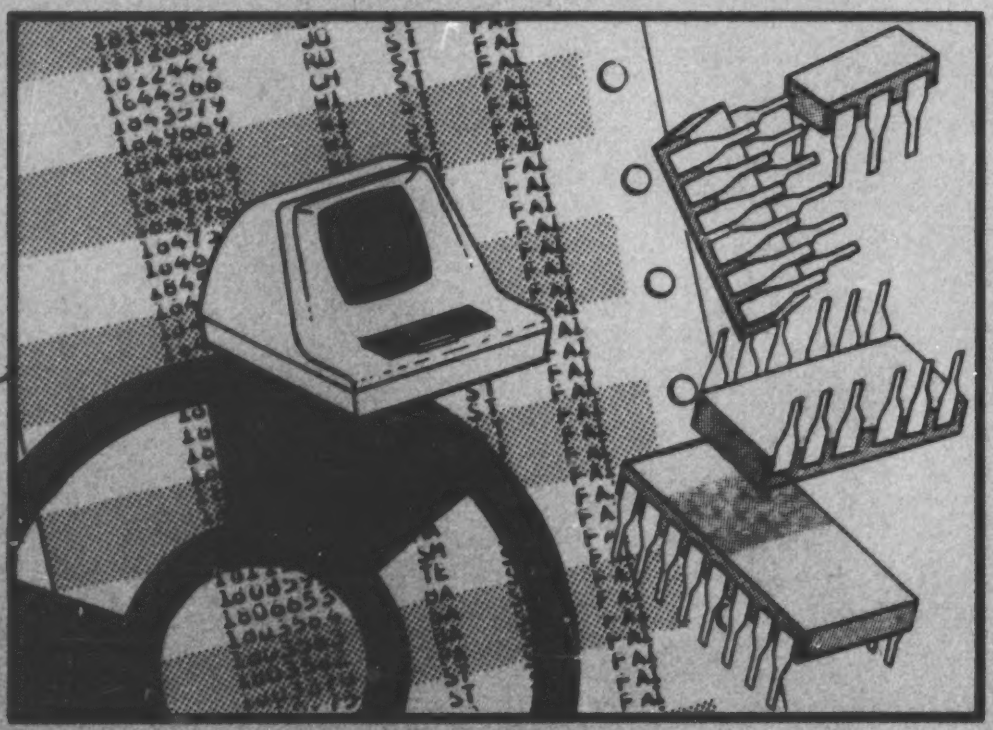
a. S-100 b. MULTIBUS c. LSI-11 d. STD

____ The system with the most peripherals available today.
____ Developed by the Digital Equipment Corporation.
____ Two systems designed primarily for 8-bit microprocessors.
____ Developed by Altair for the 8080 microprocessor.
____ Has 16 lines for data bus.
____ Developed by Pro-Log, has the fewest connector pins.
____ Will probably be the bus system interface most widely used.
____ Most experts feel that the two dominant busing systems of the future will be: _____ and _____.

ENERGY TECHNOLOGY

CONSERVATION AND USE

MICROCOMPUTER HARDWARE



MODULE MH-07

TROUBLESHOOTING
MICROCOMPUTER
COMPONENTS

TECHNICAL EDUCATION RESEARCH CENTER - SOUTHWEST
4800 LAKEWOOD DRIVE, SUITE 5
WACO, TEXAS 76710

INTRODUCTION

This module, MH-07, "Troubleshooting Microcomputer Components," provides a starting point for trying to understand what is wrong when a microcomputer does not work.

PREREQUISITES

The student should have completed Modules MH-01 through MH-06 of the course, Microcomputer Hardware.

OBJECTIVES

Upon completion of this module, the student should be able to:

1. Identify microcomputer integrated circuits (ICs) and determine their logic and function from manuals.
2. Locate the source of static hardware difficulties in small microcomputer systems.
3. Identify the functions of test equipment, such as logic analyzers, as this equipment is applied to the dynamic testing of microcomputer equipment.
4. Identify the following terms:
   a. Troubleshooting
   b. Self-diagnostics
   c. Logic probe
   d. Logic pulser
   e. Current tracer
   f. Logic clip
   g. Logic analyzer
   h. Logic comparator

i. Program tracing
j. Single-stepping

SUBJECT MATTER

In previous modules, the student has been exposed to the various components that make up a microcomputer system, as well as to how these components work together to achive specific goals.

One of the primary benefits of this exposure will be in identifying and repairing a microcomputer malfunction.

In order to design and manufacture electronic devices, humans have created complex machinery to expand their senses into the world of very small and very fast electrical phenomenon.

Microcomputers have become comprised of fewer and fewer discrete components and more and more integrated circuitry. This means that the exact nature of what is happening inside a microcomputer has become increasingly invisible to the human observer. The use of test instruments as translators of these electrical events into abstractions that are meaningful to humans enables detailed analysis. This analysis can pinpoint the causes of malfunctions. This process is called troubleshooting.

The instruments used to analyze microcomputers are sometimes more complex than the microcomputer itself. However, the use of these instruments can greatly improve the speed and effectiveness of microcomputer repair.

Troubleshooting is not the same as debugging. The term "troubleshooting" generally assumes that the microcomputer has at one time worked properly, in terms of both hardware and software. Debugging involves working out problems in the software.

The most important aspect of troubleshooting involves distinguishing between operating failures caused by programming errors (or software) and those problems caused by hardware failure.

Troubleshooting microprocessors and microcomputers will become easier as technology moves further away from discrete components toward modular replacement and self-diagnostics.

More and more microcomputer hardware is being designed with these self-diagnostic features. Self-diagnostic features include programs that perform internal circuit analysis that aids in pinpointing computer malfunction. Some diagnostic programs are so sophisticated they can pinpoint a problem to a specific integrated circuit chip on a particular board.

Even today, the circuits contained within an integrated circuit package are not accessible for troubleshooting. If the device fails to operate properly, it is simply replaced. Replacement does, however, require a knowledge of IC pin designations and internal functions.

It is possible to enlist the microcomputer's help in identifying a problem. This capability, coupled with the trend toward modular replacement, indicates that the student will probably need to do no more than isolate a specific modular component or IC board containing the source of the problem. The faulty component can then be replaced with an identical component from an on-site spares kit.

Most of today's microcomputer manufacturers offer spares kit capability. On contemporary equipment, board or component repairs can be made at regional centers; the part is then returned to the job site to become the new spare.

But for microcomputer equipment that does not have this capability, logic probes, logic clips, oscilloscopes, and logic analyzers are necessary in order to do an effective job of troubleshooting.

INTRODUCTION TO TROUBLESHOOTING EQUIPMENT

Most of the equipment used to troubleshoot microcomputers is the same equipment that is used to repair or analyze any digital electronic equipment.

Above all, the student should read and be familiar with the instruction manuals governing any test equipment – as well as manuals that deal with the microcomputer itself.

Most of the voltages within a microcomputer are less than 5 volts. However, the 115-volt line input to the main power supply – or the 9,000-volt anode voltage for the cathode-ray tube (CRT) used in many computer terminals – carries sufficient electricity to cause death if improperly handled.

One must use only shielded, high-voltage, television service-type probes when measuring voltages associated with CRTs.

One must connect the ground of all test instruments to the ground of the microcomputer being serviced. Use of an isolation transformer is recommended if there is a chance that the microcomputer ground is connected to either side of the power line – or to any electrical potential above ground.

In a defective microcomputer – or any electrical device for that matter – high voltages may appear in unexpected places.

One must not forget to bleed charged capacitors. Care should be taken that test leads are not placed where they can short out portions of the microcomputer circuitry. It is best to connect test leads to high voltage points without power to the microcomputer. One must also check the test leads for damaged insulation that might cause shorts. Test leads should be disconnected as soon as the test has been completed.

When working with high voltages, it is always preferable to have someone nearby to assist in case of an accident.

LOGIC PROBES AND PULSERS

Logic probes detect and indicate logic levels. A typical logic probe will glow dimly when not in use, or when a measured level is between the voltages required for a logic 1 and a logic 0. A logic probe will glow brightly to indicate logic 1. If the logic probe does not glow, this indicates logic 0. When measuring pulsating inputs, the probe light will flash at a slow rate for cycles up to a specified level. The probe light flashes at a fast rate for cycle rates above the specified level.

An open circuit or a high-impedance state found in three-state devices will cause the probe light to remain dimly lit.

Probes are normally used in conjunction with logic pulsers, a similar hand-held device that injects controlled pulses into digital circuitry. This enables the testing of logic gates and other ICs independently of the computer.

Both devices may be powered from the microcomputer power supply or from regulated d.c. power supplies – if the grounds from the microcomputer and the separate power supply are connected together.

CURRENT TRACER

Another hand-held probe, a current tracer, helps locate short circuits by following the path of a low-impedance fault.

The probe has an adjustable sensitivity that causes an indicator lamp to light whenever the probe detects the magnetic field generated by a pulsing current.

This probe may also be used in conjunction with the logic pulser to excite the circuit.

LOGIC CLIPS AND COMPARATORS

Logic clips are similar to multiple logic probes that clip on to ICs and continuously register the logic states of the various pins.

Logic clips require no external power source, because they automatically draw power from the power supply pin of the IC under test.

Due to the increasing variance of IC pin packages – from 16 to 24 to 40 and higher – logic clips must match the pin configuration of the IC under test to be of use.

Logic comparators are similar to logic clips, but use a reference IC that is known to be working properly to detect and indicate any variation between the reference levels and those of the IC under test.

Once again, the pin configurations must match for the device to be of value.

Similar devices are used to check entire circuit boards by comparing a new or known-good board with the board under suspicion.

Substitution of components is not advisable until a problem has been definitely located. Random substitution will cause inconsistent failure and might actually disguise the required action necessary to cure the computer malfunction.

OSCILLOSCOPES

An oscilloscope uses a cathode-ray tube display to provide a graphic representation of an electrical signal.

The sensitivity of an oscilloscope is adjustable not only to its voltage, resistance, or current characteristics, but also to its frequency and its occurrence in time.

Many oscilloscopes offer a multiple-trace capability to compare more than one signal at a time. However, the requirements of comparing the 8 or 16 simultaneous conditions present on the buses of microcomputers are better accomplished by a logic analyzer.

LOGIC ANALYZERS

Modern logic analyzers offer many capabilities that speed up the analysis of how a microcomputer is operating.

Logic analyzers use microcomputers coupled with oscilloscope-type circuitry to go to a specified location within a program and indicate the conditions of the logic circuits, buses, and so forth, at that point.

The multi-trace capabilities of logic analyzers can produce timing displays similar to those contained in most microcomputer hardware manuals for easy comparison.

These devices can also list data at specified regions in tabular form, or map conditions using special display techniques.

These analyzers can also be programmed to execute test sequences that apply pulses to specified points and then measure the results.

Many of the newer logic analyzers offer complex and useful routines to perform automatic, detailed analysis of the microcomputer circuitry under test.

TROUBLESHOOTING TECHNIQUES

There are two important categories of information that are useful in troubleshooting microcomputers. These are the instruction set introduced in the course, Microcomputer Operation, and the internal architecture of the microcomputer, as displayed in the block diagrams of the KIM-1, as shown in Module MH-06 (as well as timing charts).

Knowledge of the instruction set is necessary in order to be able to distinguish between hardware and software-initiated problems. For example, it must be known that, at the specific point within the computer program being examined, any HIGH/LOW condition on any of the address or data buses is intentional and not a mistake.

By slowing down the operation of a microcomputer to the point where it executes a single instruction upon command, these internal values can be measured and compared with a program listing.

This method of slowing down the microcomputer operation is commonly known as single-stepping, and is a form of static testing. This is the opposite of dynamic testing methods, which will be discussed later.

The comparison of measured values with a program listing is called program tracing, or simply, tracing. The clock pulses that trigger the operations can be introduced by the microcomputer by means of a single-step switch or push-button; clock pulses may also be generated by a logic pulser.

Errors or differences in measured logic levels from a program listing indicate hardware problems, since in troubleshooting it is assumed that the bugs have been worked out of the software.

Single-stepping to a point well into a long or complicated program is tedious.

The use of logic analyzers has greatly improved the ability of a troubleshooter to identify malfunctions quickly.

The goal of most troubleshooting today is to narrow the potential causes of any malfunction by means of observing any deviations from normal operating conditions.

This is the primary purpose of the diagnostic programs being included in today's microprocessors.

Once a specific area has been isolated as containing the cause of a malfunction, an increasingly important consideration in the subsequent effort to correct any problem will be the cost of replacing the failed module.

Current trends indicate that it will become increasingly less cost-efficient to troubleshoot individual components.

There are several reasons for this. The machinery used by the microcomputer manufacturer during development and testing stages enabled the manufacturer to invent very cost-effective systems for rapidly identifying and correcting problems that occur within a component.

Once a troubleshooter has isolated a problem to a specific IC (or memory board, or I/O board) replacement will become the most popular and inexpensive form of repair.

The technician should check the warranty requirements for equipment still under a warranty agreement. Many manufacturers of microcomputer equipment will not honor a warranty agreement if the customer has attempted to troubleshoot the board.

However, other manufacturers welcome on-site people who can troubleshoot to the extent of replacing defective boards from an on-site spares kit.

USING A LOGIC ANALYZER

The best method of troubleshooting a microcomputer utilizes a logic analyzer to monitor a program running at normal operating speed. This is called dynamic testing.

Not all logic analyzers manufactured today will have all of the features mentioned here. One should consult the operating manual for specific features and operating instructions.

Most logic analyzers will have a CRT-type display and a keyboard. Various probes connect the read inputs of the analyzer with the internal buses and clock lines of the microcomputer under test.

A program trace can usually be instructed to begin and end at a specified instruction and even follow selected branches within the program.

The data requested to be monitored at each program step can be formatted to be displayed in one of many specifications.

Some analyzers can be set to examine data when a specified number of clock pulses have passed from a trigger point. This feature is called digital delay.

Microcomputer timing rates can be compared through count measurement comparisons of the elapsed time, or through the number of events during execution of a specified program.

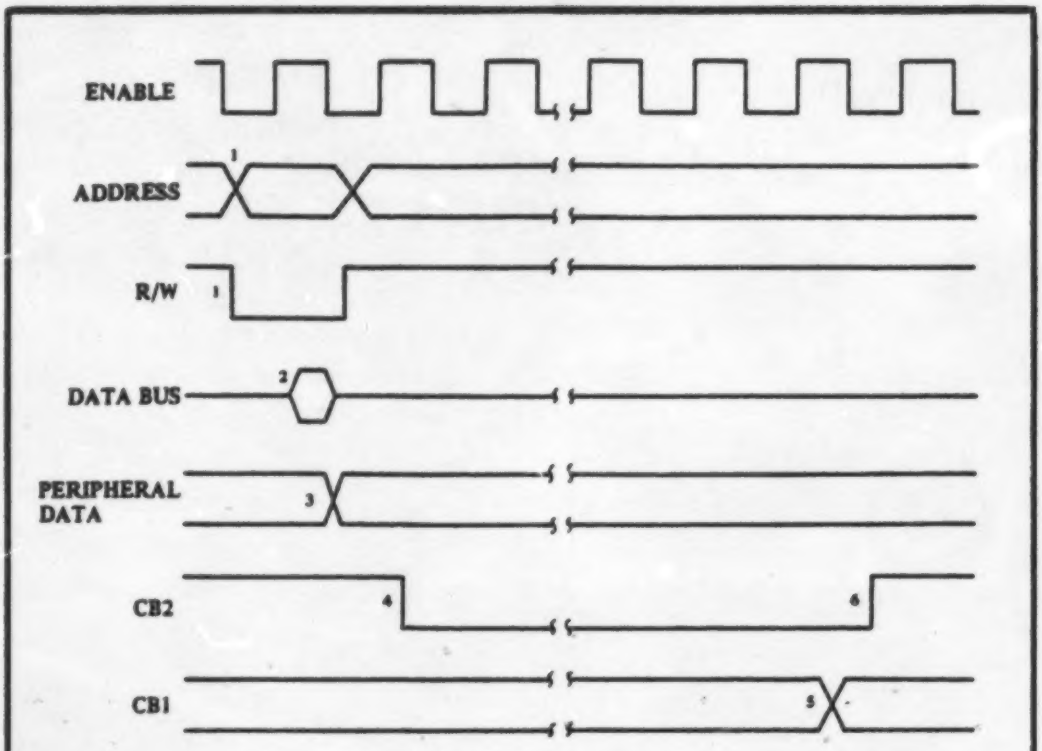
For example, investigations of this type can pinpoint problems by showing a consistent time interval between the

execution of several program steps, followed by an abnormally long time interval between other steps. This could indicate that some malfunction is occurring during the time gap and pinpoint that point for further investigation.

Graphing or mapping may be available on the logic analyzer as a means of monitoring program flow.

Most logic analyzers have the ability to sample all of the parallel lines of a bus simultaneously at a variable sampling rate, and then generate a timing diagram that graphically compares the lines. This facilitates the comparison of timing relationships within an event to manual diagrams.

Figure 1 is typical of timing diagrams found in most microcomputer hardware manuals. The particular diagram illustrated in Figure 1 shows the write handshake sequence for the KIM-1 microcomputer.



1. Processor puts out address of peripheral device and changes R/W signal to write enable (LOW).

2. During phase two, processor puts out data on Data Bus.

3. Data from the processor are accepted by the MCS62520 on the falling edge of the enable clock.

4. Peripheral interface device now begins the handshake by signaling the peripheral device that data are available to read on the output port.

5. When the external peripheral device reads the data on the output port, it will respond by a change in CB1.

6. This change in CB1 is followed by a positive transition of CB2, signaling the processor that data were accepted.

Figure 1. KIM-1 Write Handshake Sequence Timing Diagram.

HANDLING INTEGRATED CIRCUITS

At the present time, most integrated circuits encountered in microcomputer equipment are common and well-documented.

If microcomputer equipment hardware manuals do not contain enough detailed information on pin designations, internal circuitry, and so forth, this information is normally available in IC master guides. Master guides can be purchased from independent publishers or from IC manufacturers.

Great care must be taken to protect chips from random or static charges, as these charges can cause short circuits within the IC. This is especially true when removing and replacing integrated circuits.

The IC tester program in the laboratory procedure in Module MH-03 is actually a diagnostic for comparison of failed ICs to reference ICs for short-circuit detection.

The use of IC sockets in situations where ICs are frequently removed is advised.

IC timing may fluctuate within variations in power supply voltage. The effects of this variation can be deduced by varying the power supply voltages UP and DOWN while monitoring the timing of IC.

Figure 1 (the system block diagram from Module MH-06) and Figure 2 (which follows) are typical of diagrams found in microcomputer hardware manuals.

Figure 2 is a typical illustration showing IC pin designations. "A" pins are the address bus; "D" pins are the data bus. The MCS6503, MCS6504, and MCS6505 microprocessors are used in KIM microcomputers. Bar lines over the pin labels for the reset line ($\overline{RES}$), interrupt request line ($\overline{IRW}$), and the non-maskable interrupt line ($\overline{NMI}$), indicate that the normal logic level of these control lines is opposite that of the others.

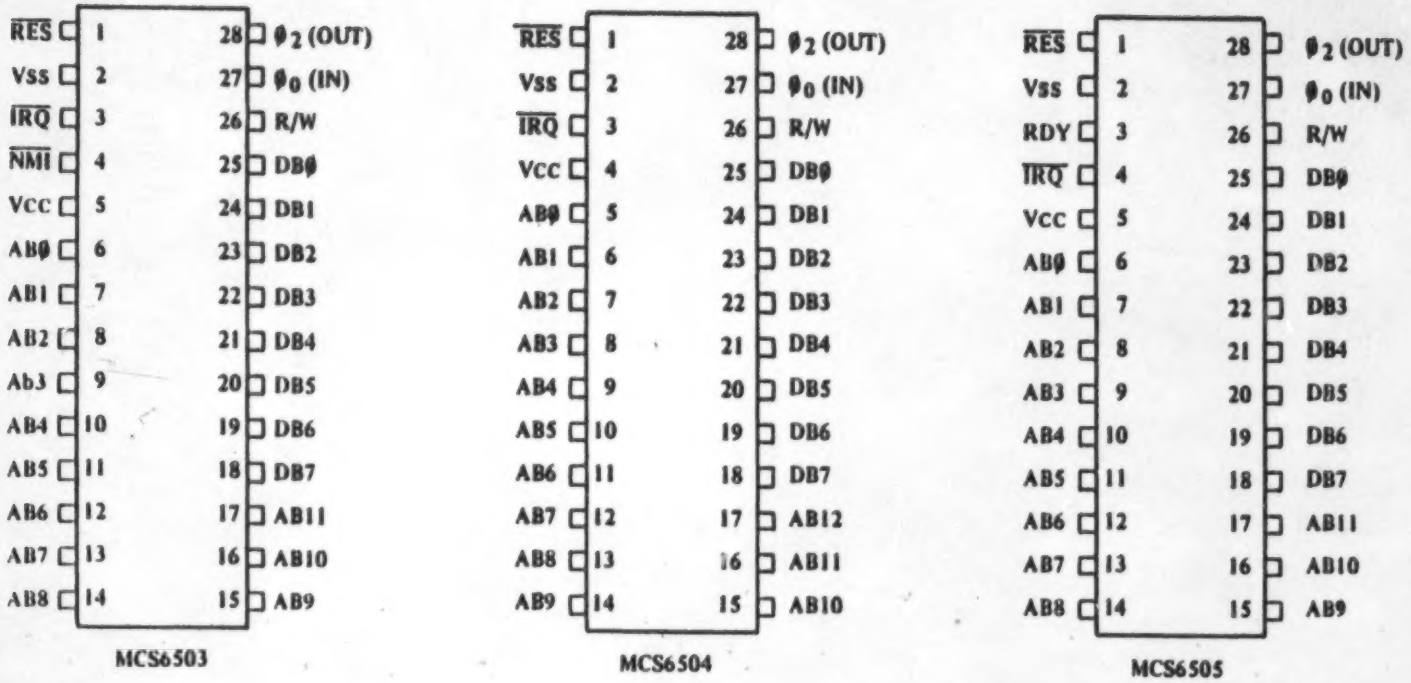


Figure 2. Diagram of IC Pin Designations.

EXERCISES

1. When using a logic probe to test TTL ICs, what should the condition of the light on the probe be when measuring the following conditions?

Condition	Light off, dim, bright, flashing, slow, flashing fast
a. +5 volts	_____
b. -2 volts	_____
c. High impedance	_____
d. 1.1 volts	_____
e. High clock speed	_____
f. No voltage	_____
g. Slow clock speed	_____

2. Which of the following program steps look suspicious?

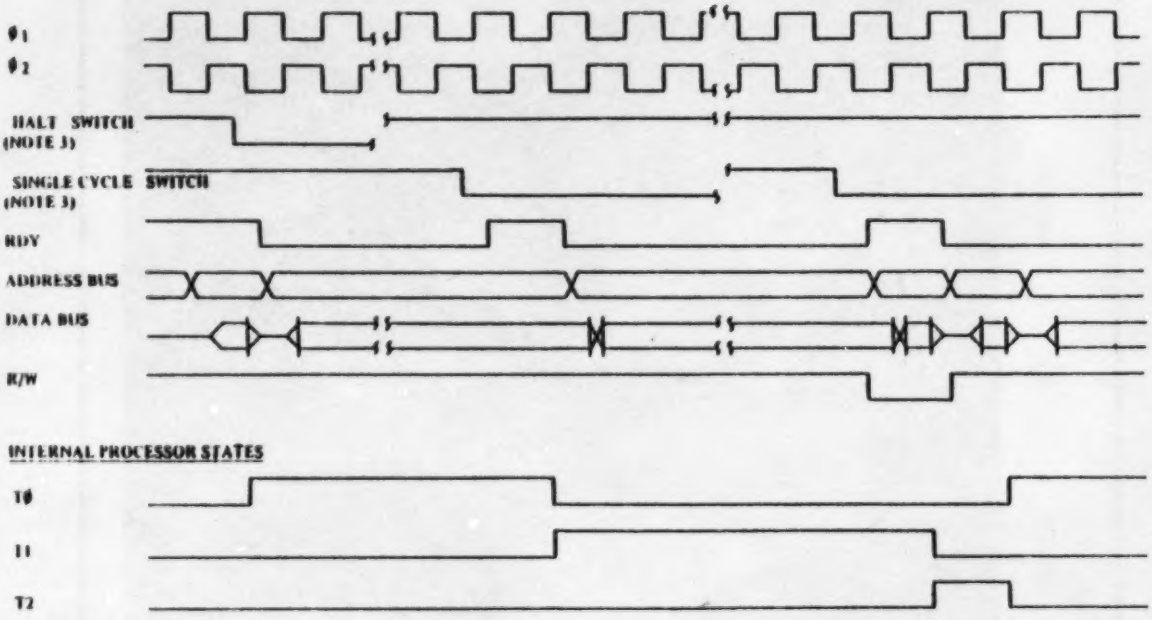
Program step	Elapsed time	Interval
Step 1	+101 microseconds	_____
Step 2	+205 microseconds	_____
Step 3	+613 microseconds	_____
Step 4	+715 microseconds	_____
Step 5	+819 microseconds	_____
Step 6	+930 microseconds	_____
Step 7	+1490 microseconds	_____
Step 8	+1600 microseconds	_____

LABORATORY MATERIALS

KIM-1 microcomputer
Logic analyzer

LABORATORY PROCEDURES

1. Use the single cycle timing diagram for the KIM-1 shown in Figure 3 for comparison with the timing display of the logic analyzer as a static test.
2. Use the address/data listings in Table 1, perform some of the dynamic tests available using the brand of logic analyzer available. Make a comparison between the results of the test and results indicated in the listings.



1. Indicates an undetermined time period during which the signal will change.

2. The data bus enters the high-impedance state during each phase one pulse. However, while the processor is stopped the data bus will appear to remain HIGH or LOW as shown.

3. Switch actuation is indicated by a LOW signal.

Figure 3. Diagram of KIM-1 Single Cycle Timing.

DATA TABLE

DATA TABLE. ADDRESS/DATA LISTINGS.

Immediate Addressing (2 cycles)

Tn	Address Bus	Data Bus	R/W	Comments
T0	PC	OP CODE	1	Fetch OP CODE
T1	PC + 1	Data	1	Fetch Data
T0	PC + 2	OP CODE	1	Next Instruction

Zero Page Addressing (3 cycles)

Tn	Address Bus	Data Bus	R/W	Comments
T0	PC	OP CODE	1	Fetch OP CODE
T1	PC + 1	ADL	1	Fetch Effective Address
T2	00, ADL	Data	1	Fetch Data
T0	PC + 2	OP CODE	1	Next Instruction

Absolute Addressing (4 cycles)

Tn	Address Bus	Data Bus	R/W	Comments
T0	PC	OP CODE	1	Fetch OP CODE
T1	PC + 1	ADL	1	Fetch LOW order Effective Address byte
T2	PC + 2	ADH	1	Fetch HIGH order Effective Address byte
T3	ADH, ADL	Data	1	Fetch Data
T0	PC + 3	OP CODE	1	Next Instruction

Indirect, X Addressing (6 cycles)

Tn	Address Bus	Data Bus	R/W	Comments
T0	PC	OP CODE	1	Fetch OP CODE
T1	PC + 1	BAL	1	Fetch Page Zero Base Address
T2	00, BAL	Data (Discarded)	1	
T3	00, BAL + X	ADL	1	Fetch LOW order byte of Effective Address
T4	00, BAL + X + 1	ADH	1	Fetch HIGH order byte of Effective Address
T5	ADH, ADL	Data	1	Fetch Data
T0	PC + 2	OP CODE	1	Next Instruction

REFERENCES

Gardner, Hershal. Handbook of Solid State Troubleshooting. Reston, VA: Reston Publishing Co., Inc., 1976.

Hall, Douglas V. Microprocessors and Digital Systems. New York: McGraw-Hill, Inc., 1980.

Lenk, John D. Handbook of Basic Electronic Troubleshooting. Englewood Cliffs, NJ: Prentice-Hall, Inc., 1977.

Lenk, John D. Handbook of Practical Microcomputer Troubleshooting. Reston, VA: Reston Publishing Co., Inc., 1979.

MOS Technology, Inc. MCS6500 Family Hardware Manual. Norristown, PA: MOS Technology, Inc., 1976.

TEST

Select the appropriate response.

1. Multiple choice

   The trend in microcomputer troubleshooting is toward ...

   a. modular replacement.
   b. random substitution.
   c. self-diagnostics.
   d. All of the above are correct.
   e. Only a and c are correct.

2. Matching

____	Logic analyzer	a.	Injects pulses
____	Logic probe	b.	Uses reference IC
____	Logic comparator	c.	Contains microprocessor
____	Logic pulser	d.	16 simultaneous readouts
____	Logic clip	e.	Off, dim, bright, flashing

3. True/False

   a. Knowledge of software has nothing to do with troubleshooting. _____
   b. If a microcomputer does not work, immediately replace all of the boards with spares that are on hand. _____
   c. As time goes by, it will become increasingly less cost-efficient to troubleshoot individual components. _____
   d. ICs are immune to static electricity. _____

4. Fill-In-the-Blank

   a. A current tracer helps to locate _____ circuits by following the path of a low-impedance fault.

b. An _____ uses a cathode-ray tube display to provide a graphic representation of an electric signal.

c. _____ is a form of static testing which slows down the operation of a microcomputer.

d. Program _____ follows the path of a program comparing measured values with a program listing.